10

```
                    GenCore version 5.1.6
             Copyright (c) 1993 - 2005  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         August 26, 2005, 03:47:51 ; Search time 211 Seconds
                                           (without alignments)
                                           7816.905 Million cell updates/sec

Title:          US-10-724-972A-2580
Perfect score:  1008
Sequence:       1 ggagtggaatcagtgagagg..........gtaaaaaagataataagtaa 1008

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters:      2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_NA:*
                1:  /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                2:  /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                3:  /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                4:  /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                5:  /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                6:  /cgn2_6/ptodata/1/ina/backfiles1.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                               SUMMARIES

                      %
Result              Query
  No.     Score     Match Length DB ID                         Description
-------------------------------------------------------------------------------
    1     222.6     22.1     668  4 US-08-956-171E-355          Sequence 355, App
    2     222.6     22.1     668  4 US-08-781-986A-355          Sequence 355, App
    3     162.2     16.1     242  4 US-08-956-171E-2556         Sequence 2556, Ap
    4     162.2     16.1     242  4 US-08-781-986A-2556         Sequence 2556, Ap
    5     158.2     15.7     183  3 US-09-134-001C-2023         Sequence 2023, Ap
    6     125.4     12.4    2115  4 US-08-956-171E-604          Sequence 604, App
    7     125.4     12.4    2115  4 US-08-781-986A-604          Sequence 604, App
    8       124     12.3     999  4 US-09-830-217-15            Sequence 15, Appl
    9       124     12.3     999  4 US-10-278-946-15            Sequence 15, Appl
   10       124     12.3    3775  4 US-08-956-171E-238          Sequence 238, App
   11       124     12.3    3775  4 US-08-781-986A-238          Sequence 238, App
   12     109.4     10.9    3046  4 US-09-710-279-3782          Sequence 3782, Ap
   13      77.2      7.7    1056  4 US-09-543-681A-2690         Sequence 2690, Ap
c  14      73.8      7.3     421  4 US-08-956-171E-383          Sequence 383, App
c  15      73.8      7.3     421  4 US-08-781-986A-383          Sequence 383, App
   16      72.2      7.2     612  4 US-09-902-540-1357          Sequence 1357, Ap
   17      70.6      7.0     975  3 US-09-071-035-37            Sequence 37, Appl
   18      70.6      7.0     984  4 US-09-134-000C-937          Sequence 937, App
   19      69.8      6.9     709  4 US-09-902-540-1583          Sequence 1583, Ap
   20      68.8      6.8     702  4 US-09-902-540-2208          Sequence 2208, Ap
c  21      68.8      6.8    7218  1 US-08-232-463-14            Sequence 14, Appl
   22      65.6      6.5     889  3 US-09-071-035-39            Sequence 39, Appl
c  23      64.6      6.4    3153  4 US-09-710-279-3348          Sequence 3348, Ap
   24      64.6      6.4    3267  4 US-09-710-279-4042          Sequence 4042, Ap
   25      64.6      6.4    3618  4 US-09-710-279-3564          Sequence 3564, Ap
   26      64.4      6.4    1044  4 US-09-710-279-461           Sequence 461, App
   27      64.4      6.4    1044  4 US-09-710-279-1267          Sequence 1267, Ap
   28      64.4      6.4    1056  3 US-09-134-001C-2655         Sequence 2655, Ap
   29      63.6      6.3    1039  4 US-09-902-540-1280          Sequence 1280, Ap
c  30      62.2      6.2    1141  4 US-09-806-708B-22           Sequence 22, Appl
   31        62      6.2     978  4 US-09-543-681A-1393         Sequence 1393, Ap
   32      61.8      6.1    9834  4 US-08-956-171E-37           Sequence 37, Appl
   33      61.8      6.1    9834  4 US-08-781-986A-37           Sequence 37, Appl
c  34      59.8      5.9    5652  4 US-09-601-198-75            Sequence 75, Appl
   35        59      5.9   19124  2 US-08-487-826B-13           Sequence 13, Appl
   36        59      5.9  640681  4 US-09-790-988-1             Sequence 1, Appli
   37      58.8      5.8     861  4 US-09-601-198-66            Sequence 66, Appl
c  38      56.8      5.6  147382  4 US-09-949-016-14624         Sequence 14624, A
   39      56.4      5.6     614  4 US-09-902-540-1318          Sequence 1318, Ap
c  40      56.4      5.6 1664976  4 US-08-916-421B-1            Sequence 1, Appli
c  41      56.4      5.6 1664976  4 US-09-692-570-1             Sequence 1, Appli
   42      55.2      5.5    4404  4 US-09-134-000C-3335         Sequence 3335, Ap
   43        55      5.5    1956  3 US-08-559-896B-1            Sequence 1, Appli
   44        55      5.5    1956  4 US-09-351-794A-1            Sequence 1, Appli
   45      54.8      5.4    1059  4 US-09-107-532A-862          Sequence 862, App


                               ALIGNMENTS


RESULT 1
US-08-956-171E-355
; Sequence 355, Application US/08956171E
; Patent No. 6593114
;  GENERAL INFORMATION:
;       APPLICANT: Charles Kunsch
;                  Gil H. Choi
;                  Patrick S. Dillon
;                  Craig A. Rosen
;                  Steven C. Barash
;                  Michael R. Fannon
;       TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;       NUMBER OF SEQUENCES: 5256
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Human Genome Sciences, Inc.
;            STREET: 9410 Key West Avenue
;            CITY: Rockville
;            STATE: Maryland
;            COUNTRY: USA
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;            COMPUTER: HP Vectra 486/33
;            OPERATING SYSTEM: MSDOS version 6.2
;            SOFTWARE: ASCII Text
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/956,171E
;            FILING DATE: 20-Oct-1997
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 60/009,861
;            FILING DATE: January 5, 1996
;            APPLICATION NUMBER: 08/781,986
;            FILING DATE: January 3, 1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Mark J. Hyman
;            REGISTRATION NUMBER: 46,789
;            REFERENCE/DOCKET NUMBER: PB248P1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (240) 314-1224
;            TELEFAX: (301) 309-8439
;  INFORMATION FOR SEQ ID NO: 355:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 668 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 355:
US-08-956-171E-355
```

```
  Query Match             22.1%;  Score 222.6;  DB 4;  Length 668;
  Best Local Similarity   74.8%;  Pred. No. 8.3e-39;
  Matches  279;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy        627 TTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAA 686
              |||  | || ||||||| | | || |||||||| |||||  |||   ||||||| |||||
Db          1 TTTATTAGCACATCCAAACTATTCATATGTTGGACAATTTTTAAACGAACTAGGATTTAA 60

Qy        687 AGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTT 746
              | | ||||||||||| ||||| || |||||||||||||| ||| | ||||||||||||||
Db         61 AAATGCATTAAGTGACGATGTAACAAAAGGTTTAAGTAAATATTTGAAAGGACCTTACTT 120

Qy        747 ACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAA 806
              |||| |  ||||||||  |||| || |  | ||||| ||||||||| |||| |||||| |
Db        121 ACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGATCATTATGACAGA 180

Qy        807 CAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAA 866
                | || |      |     ||  | | ||  |  || || ||||| |   |||||| ||
Db        181 TCATGCTAAAAAAGATTCTGCTGAATTCAAGAAGTTACAAGAAGATGCAACATGGAAAAA 240

Qy        867 ATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATC 926
               || || || || |||||| | || || |||||| | ||||||||  | |||||||||||
Db        241 GTTGAATGCAGTTAAAAATAATCGCGTGGATATTGTTGACCGTGATGTTTGGGCAAGATC 300

Qy        927 ACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGA 986
               ||||| ||||||||||| ||||||||||| |||||||||||||||||||| || |||||
Db        301 TCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTTGTTGAATTATCAAAAAAAGA 360

Qy        987 TAGTAAAAAAGAT 999
                  ||  ||| |
Db        361 ACAAAAGTAAGGT 373


RESULT 2
US-08-781-986A-355
; Sequence 355, Application US/08781986A
; Patent No. 6737248
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  5255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,986A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Bob
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PB248PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  355:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  668 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
```

```
US-08-781-986A-355

  Query Match             22.1%;  Score 222.6;  DB 4;  Length 668;
  Best Local Similarity   74.8%;  Pred. No. 8.3e-39;
  Matches  279;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy        627 TTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAA 686
              |||  | || ||||||| | | || |||||||| |||||  |||   ||||||| |||||
Db          1 TTTATTAGCACATCCAAACTATTCATATGTTGGACAATTTTTAAACGAACTAGGATTTAA 60

Qy        687 AGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTT 746
              | | ||||||||||| ||||| || |||||||||||||| ||| | ||||||||||||||
Db         61 AAATGCATTAAGTGACGATGTAACAAAAGGTTTAAGTAAATATTTGAAAGGACCTTACTT 120

Qy        747 ACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAA 806
              |||| |  ||||||||  |||| || |  | ||||| ||||||||| |||| |||||| |
Db        121 ACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGATCATTATGACAGA 180

Qy        807 CAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAA 866
                | || |      |     ||  | | ||  |  || || ||||| |   |||||| ||
Db        181 TCATGCTAAAAAAGATTCTGCTGAATTCAAGAAGTTACAAGAAGATGCAACATGGAAAAA 240

Qy        867 ATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATC 926
               || || || || |||||| | || || |||||| | ||||||||  | |||||||||||
Db        241 GTTGAATGCAGTTAAAAATAATCGCGTGGATATTGTTGACCGTGATGTTTGGGCAAGATC 300

Qy        927 ACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGA 986
               ||||| ||||||||||| ||||||||||| |||||||||||||||||||| || |||||
Db        301 TCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTTGTTGAATTATCAAAAAAAGA 360

Qy        987 TAGTAAAAAAGAT 999
                  ||  ||| |
Db        361 ACAAAAGTAAGGT 373


RESULT 3
US-08-956-171E-2556
; Sequence 2556, Application US/08956171E
; Patent No. 6593114
;   GENERAL INFORMATION:
;        APPLICANT: Charles Kunsch
;                   Gil H. Choi
;                   Patrick S. Dillon
;                   Craig A. Rosen
;                   Steven C. Barash
;                   Michael R. Fannon
;        TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;        NUMBER OF SEQUENCES: 5256
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/956,171E
;             FILING DATE: 20-Oct-1997
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/009,861
;             FILING DATE: January 5, 1996
;             APPLICATION NUMBER: 08/781,986
;             FILING DATE: January 3, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mark J. Hyman
;             REGISTRATION NUMBER: 46,789
;             REFERENCE/DOCKET NUMBER: PB248P1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (240) 314-1224
;            TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 2556:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 242 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2556:
US-08-956-171E-2556

  Query Match             16.1%;  Score 162.2;  DB 4;  Length 242;
  Best Local Similarity   79.3%;  Pred. No. 6.4e-26;
  Matches  191;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy        208 GATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAAT 267
              ||||| ||||  || ||||| |||||||| ||||| || || ||||||   || |||||
Db          2 GATGCATTAGCAGCATTAGACGTTAAACCAGTTGGTATTGCTGATGATGGTAAGAAAAAA 61

Qy        268 CGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAG 327
              ||||| ||||||||| | ||||| ||||||||  | || ||||||||||| ||||||||
Db         62 CGTATCATTAAACCAGTTAGAGAAAAAATTGGGGATTATACTTCTGTAGGTACACGTAAA 121

Qy        328 CAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAAT 387
              || || |||||||||||||| |||||| | ||||| ||||||||||| |||||||  | |
Db        122 CAGCCAAACTTAGAAGAAATTAGTAAATTAAAACCGGATTTAATTATCGCTGATAGCAGT 181

Qy        388 AGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGT 447
              ||||| | ||||||| ||||||| ||||  |||||||| || ||  |   | | || |||
Db        182 AGACATANAGGTATTAATAAAGAATTAANCAAAATTGCACCAACATTATCATTAAAGAGT 241

Qy        448 T 448
              |
Db        242 T 242


RESULT 4
US-08-781-986A-2556
; Sequence 2556, Application US/08781986A
; Patent No. 6737248
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  5255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,986A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Bob
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PB248PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  2556:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  242 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-781-986A-2556

  Query Match             16.1%;  Score 162.2;  DB 4;  Length 242;
  Best Local Similarity   79.3%;  Pred. No. 6.4e-26;
  Matches  191;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy        208 GATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAAT 267
              ||||| ||||  || ||||| |||||||| ||||| || || ||||||   || |||||
Db          2 GATGCATTAGCAGCATTAGACGTTAAACCAGTTGGTATTGCTGATGATGGTAAGAAAAAA 61

Qy        268 CGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAG 327
              ||||| ||||||||| | ||||| ||||||||  | || ||||||||||| ||||||||
Db         62 CGTATCATTAAACCAGTTAGAGAAAAAATTGGGGATTATACTTCTGTAGGTACACGTAAA 121

Qy        328 CAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAAT 387
              || || |||||||||||||| |||||| | ||||| ||||||||||| |||||||  | |
Db        122 CAGCCAAACTTAGAAGAAATTAGTAAATTAAAACCGGATTTAATTATCGCTGATAGCAGT 181

Qy        388 AGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGT 447
              ||||| | ||||||| ||||||| ||||  |||||||| || ||  |   | | || |||
Db        182 AGACATANAGGTATTAATAAAGAATTAANCAAAATTGCACCAACATTATCATTAAAGAGT 241

Qy        448 T 448
              |
Db        242 T 242


RESULT 5
US-09-134-001C-2023
; Sequence 2023, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2023
;   LENGTH: 183
;   TYPE: DNA
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2023

  Query Match             15.7%;  Score 158.2;  DB 3;  Length 183;
  Best Local Similarity   98.2%;  Pred. No. 4.3e-25;
  Matches  160;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        846 AAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGA 905
              ||||   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AAAAAGACCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGA 80

Qy        906 CCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACT 965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 CCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACT 140

Qy        966 TGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
              |||||||||||||||||||||||||||||||||||||||||||
Db        141 TGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 183


RESULT 6
US-08-956-171E-604
; Sequence 604, Application US/08956171E
; Patent No. 6593114
;   GENERAL INFORMATION:
;        APPLICANT: Charles Kunsch
;                   Gil H. Choi
;                   Patrick S. Dillon
;                   Craig A. Rosen
;                   Steven C. Barash
;                   Michael R. Fannon
;        TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;        NUMBER OF SEQUENCES: 5256
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/956,171E
;             FILING DATE: 20-Oct-1997
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/009,861
;             FILING DATE: January 5, 1996
;             APPLICATION NUMBER: 08/781,986
;             FILING DATE: January 3, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mark J. Hyman
;             REGISTRATION NUMBER: 46,789
;             REFERENCE/DOCKET NUMBER: PB248P1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (240) 314-1224
;             TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 604:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2115 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 604:
US-08-956-171E-604

  Query Match             12.4%;  Score 125.4;  DB 4;  Length 2115;
  Best Local Similarity   53.2%;  Pred. No. 8.1e-18;
  Matches  289;  Conservative    1;  Mismatches  247;  Indels    6;  Gaps    1;

Qy        458 ATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAG 517
              ||||||||| ||||||||| || ||||||||| |  | |||||  |||| |||||  |||
Db          1 ATTATAATGCAAATATTGAAGCATTTAAAACAGTCGCTAAAGCAGTAGGCAAAGAGAAAG 60

Qy        518 AAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAA 577
              ||||  | || ||  | ||| | || ||||| | | | |  ||   || |||  ||||
Db         61 AAGGCGAGAAGCGTCTGGAAAAGCATGATAAAATATTAGCGGAGATTAGAAAGAAAATTG 120

Qy        578 CTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTC 637
                  |  ||      ||   | |||  |    |      | | | ||||| || ||  |
Db        121 AACAGAGTACGTTAAAATCTGCATTTGCATTCGGTATCTCAAGAGCAGGTATGTTTATTA 180

Qy        638 ATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAA 697
              ||       |  | | | | || |||||| ||| | || | ||| || ||   | | | |
Db        181 ATAATGAAGATACATTTATGGGACAATTCTTAATTAAAATGGGTATTCAACCTGAAGTCA 240

Qy        698 GTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACA 757
                 | :|    ||||    | |  || |          || |||||  || |  | || |
Db        241 MAAAARAMAAAACTACGCATGTTGGTGAACGCAAGGGTGGTCCTTATATATATTTAAATA 300

Qy        758 CTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTT 817
               ||||    |  |  |  | |||||  |   |||| |  ||   ||  ||  |  ||
Db        301 ATGAAGAACTTGCCAATATCAATCCAAAAGTTATGATTTTAGCCACTGACGGAAAAACGG 360

Qy        818 CTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTG 877
                ||  |         ||   || |     |   | | || |||||   |||||| ||||
Db        361 ACAAAAA------TAGAACGAAATTCATTGATCCTGCAGTTTGGAAATCATTAAAAGCTG 414

Qy        878 TGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAA 937
              ||||| || |    ||| ||    | |||||  |    |||   | |||| | ||  | |
Db        415 TGAAAGATAACAAAGTTTATGACGTTGACCGAAATAAGTGGTTGAAATCAAGGGGGATTA 474

Qy        938 TTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAG 997
              |  |    |||   |||||| ||||  | |   || |  |  | ||||  |   |||||
Db        475 TCGCAAGTGAAAGTATGGCAGAAGATTTAGAAAAAATTGCAGAAAAAGCAAAATAAAAAT 534

Qy        998 ATA 1000
              | |
Db        535 ACA 537


RESULT 7
US-08-781-986A-604
; Sequence 604, Application US/08781986A
; Patent No. 6737248
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  5255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,986A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Bob
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PB248PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  604:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2115 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-781-986A-604

  Query Match             12.4%;  Score 125.4;  DB 4;  Length 2115;
  Best Local Similarity   53.2%;  Pred. No. 8.1e-18;
  Matches  289;  Conservative    1;  Mismatches  247;  Indels    6;  Gaps    1;

Qy        458 ATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAG 517
              ||||||||| ||||||||| || ||||||||| |  | |||||  |||| |||||  |||
Db          1 ATTATAATGCAAATATTGAAGCATTTAAAACAGTCGCTAAAGCAGTAGGCAAAGAGAAAG 60

Qy        518 AAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAA 577
              ||||  | || ||  | ||| | || ||||| | | | |  ||   || |||  ||||
Db         61 AAGGCGAGAAGCGTCTGGAAAAGCATGATAAAATATTAGCGGAGATTAGAAAGAAAATTG 120
```

```
Qy          578 CTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTC 637
                    |  ||      ||   | |||  |    |      | | | ||||| || ||  |
Db          121 AACAGAGTACGTTAAAATCTGCATTTGCATTCGGTATCTCAAGAGCAGGTATGTTTATTA 180

Qy          638 ATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAA 697
                ||       |  | | | | || |||||| ||| | || | ||| || ||   | | | |
Db          181 ATAATGAAGATACATTTATGGGACAATTCTTAATTAAAATGGGTATTCAACCTGAAGTCA 240

Qy          698 GTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACA 757
                   | :|    ||||    | |  || |          || |||||  || |  | || |
Db          241 MAAAARAMAAAACTACGCATGTTGGTGAACGCAAGGGTGGTCCTTATATATATTTAAATA 300

Qy          758 CTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTT 817
                 ||||    |  |  |  | |||||  |   |||| |  ||   ||  ||  |  ||
Db          301 ATGAAGAACTTGCCAATATCAATCCAAAAGTTATGATTTTAGCCACTGACGGAAAAACGG 360

Qy          818 CTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTG 877
                  ||  |         ||   || |     |   | | || |||||   |||||| ||||
Db          361 ACAAAAA------TAGAACGAAATTCATTGATCCTGCAGTTTGGAAATCATTAAAAGCTG 414

Qy          878 TGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAA 937
                ||||| || |    ||| ||    | |||||  |    |||   | |||| | ||  | |
Db          415 TGAAAGATAACAAAGTTTATGACGTTGACCGAAATAAGTGGTTGAAATCAAGGGGGATTA 474

Qy          938 TTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAG 997
                |  |    |||   |||||| ||||  | |   || |  |  | ||||  |   |||||
Db          475 TCGCAAGTGAAAGTATGGCAGAAGATTTAGAAAAAATTGCAGAAAAAGCAAAATAAAAAT 534

Qy          998 ATA 1000
                | |
Db          535 ACA 537

RESULT 8
US-09-830-217-15
; Sequence 15, Application US/09830217
; Patent No. 6521441
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc.
;  TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
;  FILE REFERENCE: PB461PCT
;  CURRENT APPLICATION NUMBER: US/09/830,217
;  CURRENT FILING DATE:  2001-04-24
;  PRIOR APPLICATION NUMBER: PCT/US99/06199
;  PRIOR FILING DATE: 1999-03-18
;  PRIOR APPLICATION NUMBER: 60/080,296
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/084,674
;  PRIOR FILING DATE: 1998-05-07
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 999
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-09-830-217-15

  Query Match             12.3%;  Score 124;  DB 4;  Length 999;
  Best Local Similarity   50.7%;  Pred. No. 1.4e-17;
  Matches  412;  Conservative    0;  Mismatches  380;  Indels   20;  Gaps    4;

Qy           27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
                ||||  ||   |  |||      |   | |  || |||     |||| |||| |    ||
Db           18 AATTAAAATGCTTGTTGTTACGCTTGCTTTCCTACTTGTTTTAGCAGGATGTAGTGGGAA 77

Qy           87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
                |      |   | |||  | |  |   |   || ||   |  |  ||| || || | |
Db           78 TTCAAATAAACAATCATCTGATAACAAAGATAAGGAAACAACTTCAATTAAACATGCAAT 137

Qy          147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
                 ||||| ||  || |        | || || |||||||||   ||  |  |
Db          138 GGGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCAC 197
```

```
Qy          207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
                ||| |     ||  |||||| |||||||||||| ||    |  ||       | | ||||
Db          198 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAA 257

Qy          267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
                 |  |   |  ||    ||| | || |  |   | |    |    ||||||   |
Db          258 ACCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACC 317

Qy          327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                   |||||||||||| ||||||  |||| | ||||| || |||||| | ||    ||
Db          318 TGCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGT 377

Qy          387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
                |||| |  ||    ||||  |  | |||  |||||| || ||             || ||
Db          378 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAG 424

Qy          447 TTTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAG 505
                |||| |  || |   |  | ||| | |  ||| |   |||   |||    ||||||||||
Db          425 TTTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAG 484

Qy          506 GTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATA 565
                | |||||| |||||| | || |     |  || |  ||||| | ||| | |  | ||
Db          485 GGAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCC 544

Qy          566 AAAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAAT 622
                |||||||   || | | | | ||||| ||  |  |  ||||   ||   || || ||| |
Db          545 AAAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACT 604

Qy          623 CAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
                   ||     || |||| |||| |  | |  || |||  |              ||  |
Db          605 TCCGT---GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 661

Qy          683 TTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTT 742
                 |  || |     | || || | |      ||     |  | |  ||  |||||    |
Db          662 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 721

Qy          743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGA 802
                 |  ||  |     |  ||||   |  |   | ||||  |||| | ||| ||  || | |
Db          722 TCCAACTTACATCTAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAA 781

Qy          803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
                 |    |  ||| | | || ||  || || ||
Db          782 AATCAGATCCAAATGCGAAAGATGCTGCATTA 813

RESULT 9
US-10-278-946-15
; Sequence 15, Application US/10278946
; Patent No. 6821754
; GENERAL INFORMATION:
;  APPLICANT: Simpson et al.
;  TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
;  FILE REFERENCE: PB461USD1
;  CURRENT APPLICATION NUMBER: US/10/278,946
;  CURRENT FILING DATE:  2002-10-24
;  PRIOR APPLICATION NUMBER: US 09/830,217
;  PRIOR FILING DATE: 2001-04-24
;  PRIOR APPLICATION NUMBER: PCT/US99/06199
;  PRIOR FILING DATE: 1999-03-18
;  PRIOR APPLICATION NUMBER: US 60/078,682
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: US 60/080,296
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: US 60/084,674
;  PRIOR FILING DATE: 1998-05-07
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
;   LENGTH: 999
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-10-278-946-15
```

```
  Query Match             12.3%;  Score 124;  DB 4;  Length 999;
  Best Local Similarity   50.7%;  Pred. No. 1.4e-17;
  Matches  412;  Conservative    0;  Mismatches  380;  Indels   20;  Gaps    4;

Qy          27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
               ||||  ||   |  |||      |   | |  || |||     |||| |||| |    ||
Db          18 AATTAAAATGCTTGTTGTTACGCTTGCTTTCCTACTTGTTTTAGCAGGATGTAGTGGGAA 77

Qy          87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
               |      |   | |||  | |  |   |   || ||   |  |  ||| || ||  | |
Db          78 TTCAAATAAACAATCATCTGATAACAAAGATAAGGAAACAACTTCAATTAAACATGCAAT 137

Qy         147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
                ||||| ||  ||  |       | || || |||||||||   ||  |  |
Db         138 GGGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCAC 197

Qy         207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
               ||| |     ||  |||||| |||||||||||| ||    |  ||       | | ||||
Db         198 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAA 257

Qy         267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
                |  |   |  ||    ||| | || |  |   | |    |    ||||||   |
Db         258 ACCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACC 317

Qy         327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                  |||||||||||| ||||||  |||| | ||||| || |||||| | ||    ||
Db         318 TGCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGT 377

Qy         387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
               |||| |  ||    ||||  |  | |||  |||||| || ||             || ||
Db         378 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAG 424

Qy         447 TTTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAG 505
               |||| |  || |   |  | ||| | |  ||| |   |||   |||    ||||||||||
Db         425 TTTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAG 484

Qy         506 GTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATA 565
               | |||||| |||||| | || |     |  || |  ||||| | ||| | |  | ||
Db         485 GGAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCC 544

Qy         566 AAAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAAT 622
               |||||||   || | | | | ||||| ||  |  |  ||||   ||   || || ||| |
Db         545 AAAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACT 604

Qy         623 CAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
                  ||     || |||| |||| |  | |  || |||  |              ||  |
Db         605 TCCGT---GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 661

Qy         683 TTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTT 742
                |  || |     | || || | |      ||     |  | |  ||  |||||    |
Db         662 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 721

Qy         743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGA 802
                |  ||  |     |  |||   |  |   | ||||  |||| | ||| ||  || | |
Db         722 TCCAACTTACATCTAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAA 781

Qy         803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
                |    |  ||| | | || ||  || || ||
Db         782 AATCAGATCCAAATGCGAAAGATGCTGCATTA 813


RESULT 10
US-08-956-171E-238
; Sequence 238, Application US/08956171E
; Patent No. 6593114
;   GENERAL INFORMATION:
;         APPLICANT: Charles Kunsch
;                    Gil H. Choi
;                    Patrick S. Dillon
;                    Craig A. Rosen
;                    Steven C. Barash
;                    Michael R. Fannon
;         TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;         NUMBER OF SEQUENCES: 5256
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Human Genome Sciences, Inc.
;              STREET: 9410 Key West Avenue
;              CITY: Rockville
;              STATE: Maryland
;              COUNTRY: USA
;              ZIP: 20850
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;              COMPUTER: HP Vectra 486/33
;              OPERATING SYSTEM: MSDOS version 6.2
;              SOFTWARE: ASCII Text
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/08/956,171E
;              FILING DATE: 20-Oct-1997
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 60/009,861
;              FILING DATE: January 5, 1996
;              APPLICATION NUMBER: 08/781,986
;              FILING DATE: January 3, 1997
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Mark J. Hyman
;              REGISTRATION NUMBER: 46,789
;              REFERENCE/DOCKET NUMBER: PB248P1
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (240) 314-1224
;              TELEFAX: (301) 309-8439
;    INFORMATION FOR SEQ ID NO: 238:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 3775 base pairs
;              TYPE: nucleic acid
;              STRANDEDNESS: double
;              TOPOLOGY: linear
;         SEQUENCE DESCRIPTION: SEQ ID NO: 238:
US-08-956-171E-238

  Query Match             12.3%;  Score 124;  DB 4;  Length 3775;
  Best Local Similarity   50.7%;  Pred. No. 1.8e-17;
  Matches  412;  Conservative    0;  Mismatches  380;  Indels   20;  Gaps    4;

Qy          27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
               ||||  ||   |  |||      |   | |  || |||     |||| |||| |    ||
Db          60 AATTAAAATGCTTGTTGTTACGCTTGCTTTCCTACTTGTTTTAGCAGGATGTAGTGGGAA 119

Qy          87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
               |      |   | |||  | |  |   |   || ||   |  |  ||| || ||  | |
Db         120 TTCAAATAAACAATCATCTGATAACAAAGATAAGGAAACAACTTCAATTAAACATGCAAT 179

Qy         147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
                ||||| ||  ||  |       | || || |||||||||   ||  |  |
Db         180 GGGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCAC 239

Qy         207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
               ||| |     ||  |||||| |||||||||||| ||    |  ||       | | ||||
Db         240 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAA 299

Qy         267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
                |  |   |  ||    ||| | || |  |   | |    |    ||||||   |
Db         300 ACCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACC 359

Qy         327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                  |||||||||||| ||||||  |||| | ||||| || |||||| | ||    ||
Db         360 TGCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGT 419

Qy         387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
               |||| |  ||    ||||  |  | |||  |||||| || ||             || ||
Db         420 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAG 466

Qy         447 TTTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAG 505
               |||| |  || |   |  | ||| | |  ||| |   |||   |||    ||||||||||
Db         467 TTTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAG 526
```

```
Qy          506 GTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATA 565
                | |||||| |||||| | || |     |  || |  ||||| | ||| | |  | ||
Db          527 GGAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCC 586

Qy          566 AAAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAAT 622
                |||||||   || | | | | ||||| ||  |  |  ||||   ||   || || ||| |
Db          587 AAAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACT 646

Qy          623 CAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
                   ||     || |||| |||| |  | |  || |||  |              ||  |
Db          647 TCCGT---GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 703

Qy          683 TTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTT 742
                 |  || |     | || || | |      ||     |  | |  ||  |||||    |
Db          704 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 763

Qy          743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGA 802
                 |  ||  |     |  ||||   |  |   | ||||  |||| | ||| ||  || | |
Db          764 TCCAACTTACATCTAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAA 823

Qy          803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
                 |    |  ||| | | || ||  || || ||
Db          824 AATCAGATCCAAATGCGAAAGATGCTGCATTA 855
```

```
RESULT 11
US-08-781-986A-238
; Sequence 238, Application US/08781986A
; Patent No. 6737248
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  5255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,986A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Bob
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PB248PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  238:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3775 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-781-986A-238

  Query Match             12.3%;  Score 124;  DB 4;  Length 3775;
  Best Local Similarity   50.7%;  Pred. No. 1.8e-17;
  Matches  412;  Conservative    0;  Mismatches  380;  Indels   20;  Gaps    4;

Qy           27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
                ||||  ||   |  |||      |   | |  || |||     |||| |||| |    ||
Db           60 AATTAAAATGCTTGTTGTTACGCTTGCTTTCCTACTTGTTTTAGCAGGATGTAGTGGGAA 119

Qy           87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
                |      |   | |||  | |  |   |   || ||   |  |  ||| || || | |
Db          120 TTCAAATAAACAATCATCTGATAACAAAGATAAGGAAACAACTTCAATTAAACATGCAAT 179

Qy          147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
                 ||||| ||  || |        | || || |||||||||   ||  |  |
Db          180 GGGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCAC 239

Qy          207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
                ||| |     ||  |||||| |||||||||||| ||    |  ||       | | ||||
Db          240 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAA 299

Qy          267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
                 |  |   |  ||    ||| | || |  |   | |    |     ||||||   |
Db          300 ACCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACC 359

Qy          327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                   |||||||||||| ||||||  |||| | ||||| || |||||| | ||    ||
Db          360 TGCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGT 419

Qy          387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
                |||| |  ||    ||||  |  | |||  |||||| || ||             || ||
Db          420 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAG 466

Qy          447 TTTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAG 505
                |||| |  || |   |  | ||| | |  ||| |   |||   |||    ||||||||||
Db          467 TTTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAG 526

Qy          506 GTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATA 565
                | |||||| |||||| | || |     |  || |  ||||| | ||| | |  | ||
Db          527 GGAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCC 586

Qy          566 AAAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAAT 622
                |||||||   || | | | | ||||| ||  |  |  ||||   ||   || || ||| |
Db          587 AAAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACT 646

Qy          623 CAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
                   ||     || |||| |||| |  | |  || |||  |              ||  |
Db          647 TCCGT---GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 703

Qy          683 TTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTT 742
                 |  || |     | || || | |      ||     |  | |  ||  |||||    |
Db          704 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 763

Qy          743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGA 802
                 |  ||  |     |  ||||   |  |   | ||||  |||| | ||| ||  || | |
Db          764 TCCAACTTACATCTAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAA 823

Qy          803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
                 |    |  ||| | | || ||  || || ||
Db          824 AATCAGATCCAAATGCGAAAGATGCTGCATTA 855
```

```
RESULT 12
US-09-710-279-3782
; Sequence 3782, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
;  APPLICANT: KIMMERLY, WILLIAM JOHN
;  TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;  FILE REFERENCE: PU3480US
;  CURRENT APPLICATION NUMBER: US/09/710,279
;  CURRENT FILING DATE:  2000-11-09
;  PRIOR APPLICATION NUMBER: 60/164,258
;  PRIOR FILING DATE: 1999-11-09
;  NUMBER OF SEQ ID NOS: 4472
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3782
;   LENGTH: 3046
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3782

  Query Match             10.9%;  Score 109.4;  DB 4;  Length 3046;
  Best Local Similarity   95.0%;  Pred. No. 2.4e-14;
  Matches  113;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy           1 GGAGTGGAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTA 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2705 GGAGTGGAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTA 2764

Qy          61 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAA 119
               ||||||||||||||||||||||||||||||||||||||||||||||| |    || |||
Db        2765 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAATGTAATATAAAA 2823


RESULT 13
US-09-543-681A-2690
; Sequence 2690, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
;  APPLICANT: GARY BRETON
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIL
;  TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 2709.1002-001
;  CURRENT APPLICATION NUMBER: US/09/543,681A
;  CURRENT FILING DATE:  2000-04-05
;  PRIOR APPLICATION NUMBER: US 60/128,706
;  PRIOR FILING DATE: 1999-04-09
;  NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 2690
;   LENGTH: 1056
;   TYPE: DNA
;   ORGANISM: Proteus mirabilis
US-09-543-681A-2690

  Query Match              7.7%;  Score 77.2;  DB 4;  Length 1056;
  Best Local Similarity   48.7%;  Pred. No. 1.6e-07;
  Matches  270;  Conservative    0;  Mismatches  278;  Indels    6;  Gaps    2;

Qy          24 AAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAA 83
               ||||   |||||   | |    ||||||| |   ||   | |  |  ||||  ||||  ||
Db         105 AAAATCATTAAGCCCACTGTTCTTATTGCTGTCTTCTCTGGTTATCGCAGGTTGTGATAA 164

Qy          84 TAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGA 143
                         |      |||   ||    ||   ||| |     |    ||| | || | 
Db         165 CGCACAAGATACCTCAACAACGCAATCCACAGAGAAACAAACTCTCACTATCGAACATGC 224

Qy         144 AGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATT 203
                 |||| || ||  |  | |||   |||||  ||   || |||||  |  |      |  
Db         225 TCAAGGCACCACTGAGATCCCTGCTCACCCACAAAAAGTGGTTGTGATGAACATGGAAAC 284

Qy         204 TGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAA 263
                 |||||    | | |||| | |  ||      ||||||  || |  |    |||    |
Db         285 ACTTGATATTGTTGATGCTCTTGGCGTACCTGTTGTTGGCCTACCACAAACAAACGTCCA 344

Qy         264 AAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACG 323
                   | ||   ||  |     |      ||        | ||||     | |||    | 
Db         345 CCTACCTAAATTTCTAGAGAAATACACCAACGAAAATGAGTACATCAACGAAGGTGGCCT 404

Qy         324 TAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAA 383
                    |||| ||||  ||| || ||||||    |   || ||| | ||| |  | | || 
Db         405 GTTTGAACCAAACTACGAAAAACTCAGTACTACTGCCCCTGATCTGATTTTAACGGGTAG 464

Qy         384 TAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAA 443
               |  |   | |   | |   |||    | |||| | |||| || ||  | |||    || |
Db         465 TCGTGCTCGC---GATGCATATGCCAAATTAAGTGAAATCGCACCATCCATTTCTATGGA 521

Qy         444 AAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCA---AAAGC 500
                |    |     |   | || ||| |   |   ||    || ||| | || |    ||  
Db         522 TATCGACAGCACACGCTTTATTGACAGCCTCACTGAGCGTACAACGACTTTAGGTCAAAT 581

Qy         501 TTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGA 560
               ||| ||||||||||||| ||| |||||||    ||||  ||    | |||  ||||||||   
Db         582 TTTTGGTAAAGAAGAACAAGCTAAAAAATTATTAGCTGACTTCAATAGCAAAATTGATAC 641

Qy         561 ATATAAAAAAGAAA 574
                   |||  | |||
Db         642 GGTAAAAGCAAAAA 655


RESULT 14
US-08-956-171E-383/c
; Sequence 383, Application US/08956171E
; Patent No. 6593114
;  GENERAL INFORMATION:
;       APPLICANT: Charles Kunsch
;                  Gil H. Choi
;                  Patrick S. Dillon
;                  Craig A. Rosen
;                  Steven C. Barash
;                  Michael R. Fannon
;       TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;       NUMBER OF SEQUENCES: 5256
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Human Genome Sciences, Inc.
;            STREET: 9410 Key West Avenue
;            CITY: Rockville
;            STATE: Maryland
;            COUNTRY: USA
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;            COMPUTER: HP Vectra 486/33
;            OPERATING SYSTEM: MSDOS version 6.2
;            SOFTWARE: ASCII Text
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/956,171E
;            FILING DATE: 20-Oct-1997
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 60/009,861
;            FILING DATE: January 5, 1996
;            APPLICATION NUMBER: 08/781,986
;            FILING DATE: January 3, 1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Mark J. Hyman
;            REGISTRATION NUMBER: 46,789
;            REFERENCE/DOCKET NUMBER: PB248P1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (240) 314-1224
;            TELEFAX: (301) 309-8439
;  INFORMATION FOR SEQ ID NO: 383:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 421 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 383:
US-08-956-171E-383

  Query Match              7.3%;  Score 73.8;  DB 4;  Length 421;
  Best Local Similarity   61.9%;  Pred. No. 7.1e-07;
  Matches  117;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy           7 GAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCA 66
               | | |  |||||||| |||||| ||| ||| |  | || ||| |  |||  ||| ||  |
Db         205 GGAACGATGAGAGGTCTAAAAACTTTTAGTATATTGGGATTAATAGTTGCCTTACTTTTA 146

Qy          67 ACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGA 126
                 |||||| ||||| ||||    |      |    ||   ||  | |  | ||||||   
Db         145 GTTGCAGCTTGTGGTAATACGGATAATTCAAGTAAAAAAGAATCATCAACTAAAGATACT 86
```

```
Qy        127 GTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTT 186
               ||    | ||  | ||| | |||||    |||||||||||| |  | |||||| | |||
Db         85 ATTTCGGTAAAAGATGAAAATGGTACAGTAAAAGTACCTAAAGATGCAAAACGTATCGTT 26

Qy        187 GTTCTTGAG 195
              ||  | |||
Db         25 GTATTAGAG 17

RESULT 15
US-08-781-986A-383/c
; Sequence 383, Application US/08781986A
; Patent No. 6737248
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  5255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,986A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Bob
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PB248PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  383:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  421 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-781-986A-383

  Query Match             7.3%;  Score 73.8;  DB 4;  Length 421;
  Best Local Similarity   61.9%;  Pred. No. 7.1e-07;
  Matches  117;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy          7 GAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCA 66
              | | |  |||||||| |||||| ||| ||| || | || ||| |  |||  ||| ||  |
Db        205 GGAACGATGAGAGGTCTAAAAACTTTTAGTATATTGGGATTAATAGTTGCCTTACTTTTA 146

Qy         67 ACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGA 126
                |||||| ||||| ||||    |      |    ||   ||   ||| | ||||||
Db        145 GTTGCAGCTTGTGGTAATACGGATAATTCAAGTAAAAAAGAATCATCAACTAAAGATACT 86

Qy        127 GTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTT 186
               ||    | ||  | ||| | |||||    |||||||||||| |  | |||||| | |||
Db         85 ATTTCGGTAAAAGATGAAAATGGTACAGTAAAAGTACCTAAAGATGCAAAACGTATCGTT 26

Qy        187 GTTCTTGAG 195
              ||  | |||
Db         25 GTATTAGAG 17

RESULT 16
US-09-902-540-1357
; Sequence 1357, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1357
;   LENGTH: 612
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(612)
;   OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357

  Query Match             7.2%;  Score 72.2;  DB 4;  Length 612;
  Best Local Similarity   50.4%;  Pred. No. 1.7e-06;
  Matches  173;  Conservative    0;  Mismatches  170;  Indels    0;  Gaps    0;

Qy        254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
              |||| ||||||||    |  |  |||  |  || | | ||||    |||| | |
Db        150 ATAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAA 209

Qy        314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
               |  || |   ||  |    ||   | || |||  | ||||   |||  | |    |  |
Db        210 AAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAATAAAATAAAA 269

Qy        374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
               |    | ||||| | |   |||   | ||| ||| |   |||  ||| |      |  |
Db        270 ATAAAAAAAATAAAATATTAAAAAAAAATTAAAAAAAAAAAAANTAAAATAAAAAAAAAA 329

Qy        434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                 || | ||||  |   |   | || |  |  |||| | | |       |||| ||  |
Db        330 AAAAANTAAAAAAATAAAAAAAAAATAAATAAAAAAAAAATAAAAAAAAAAAAATAAAAT 389

Qy        494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
                |||  |  |   ||| || || ||   ||||||     | || || |    || ||||
Db        390 ATAAAAATAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAATTAAAAAAAA 449

Qy        554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAA 596
               | || ||   | |||| ||| ||      | |||||  ||||
Db        450 ATTAAAAAATAATAAAAAAAAAAAAATAAAAAAAAAAAAAAAA 492

RESULT 17
US-09-071-035-37
; Sequence 37, Application US/09071035
; Patent No. 6448043
;  GENERAL INFORMATION:
;    APPLICANT:  Gil H. Choi
;    TITLE OF INVENTION:  Enterococcus faecalis Polynucleotides and Polypeptides
;    NUMBER OF SEQUENCES:  496
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/071,035
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  A. Anders Brookes
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB369P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  37:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  975 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-09-071-035-37

  Query Match             7.0%;  Score 70.6;  DB 3;  Length 975;
  Best Local Similarity   48.8%;  Pred. No. 4.1e-06;
  Matches  317;  Conservative    0;  Mismatches  314;  Indels   18;  Gaps    4;

Qy          48 ATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAA 107
               | ||||  | |||| |||   |  | ||  |  ||| |          |    |||   |
Db          66 ACTGTTAATGTTAAGTGCTTGTCAAACAAATAAAAAAACAGCAGATTCTGCAACAACAGA 125

Qy         108 AGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAA 167
               |    || | |||   | |||   | ||||  ||    | |||      |  || || ||
Db         126 AACAACAGCTAAAACGGAAGTCACAGTCAAAGACACCAATGGTCAATTAACCGTTCCCAA 185

Qy         168 ACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGA 227
               | | |||||    || |||||| ||||  ||   |   | ||| |  | | |||  |||
Db         186 AAATCCTAAGAAAGTCGTTGTTTTTGATAATGGTTCCTTGGATACAATGGATGCACTAGG 245

Qy         228 TGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAG 287
               |||          | | |||  |     |  ||||| |  | |     |     |  |
Db         246 TGTCGGTGACCGCGTGGTAGGTGCGCCAACTAAAAATATCCCTGCGTATTTGAAAAAATA 305

Qy         288 AGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAAT 347
                 | ||| ||| |         || | |||      |||  ||||  | |||||| ||||
Db         306 CCAAAAAGTTGAA---------TCAGCAGGCGGCATTAAAGAACCAGATTTAGAAAAAAT 356

Qy         348 CAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAA 407
               || | |||| |||||||| |||||||||  |  |  |  |  ||   ||| |  | |  |
Db         357 CAATCAACTAAAACCAGACTTAATTATTATTTCTGGTCGTCAAC---AAGATTATCAAGA 413

Qy         408 AGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGA 467
               | | |||||    ||||| || || ||| |  |          |    | ||  |   |
Db         414 ACAATTAAAAGCCATTGCGCCAACCATTTACTTAGCTGTAGATGCCAAAAATCCTTGGGC 473

Qy         468 ---AAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAA 524
                  ||  |   |   | |  ||||  |    |    ||| | ||||||||| | || |||
Db         474 ATCAACGAAACAAAATATCGAAACGTTAGGCACTATTTTTGATAAAGAAGAGGTAGCTAA 533

Qy         525 AAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGA 584
               | ||    ||   |     || ||  ||||||  ||    |||||| |   | |    |
Db         534 AGAAAAAATAACTGGCTTAGAAAAAGAAATTGCTGACGTGAAAAAACA---AGCAGAAGC 590

Qy         585 TAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAG 644
               ||      | || |   |   ||   ||| |    ||  |    |    ||| |   ||
Db         591 TAGCGCGAATAATGCGCTTGTTGTGTTAGTTAACGAAGGACAACTTTCCGCTTACGGAAA 650

Qy         645 CAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCA 693
                  ||||  | | |||  | | |    |  | | || || |||| ||||
Db         651 AGGCTCTCGTTTCGGTTTAATTCATGATACATTTGGCTTCAAAGCAGCA 699


RESULT 18
US-09-134-000C-937
; Sequence 937, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;  TITLE OF INVENTION:  ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 032796-032
;  CURRENT APPLICATION NUMBER: US/09/134,000C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/055,778
;  PRIOR FILING DATE: 1997-08-15
;  NUMBER OF SEQ ID NOS: 6812
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 937
;   LENGTH: 984
;   TYPE: DNA
;   ORGANISM: Enterococcus faecalis
US-09-134-000C-937

  Query Match             7.0%;  Score 70.6;  DB 4;  Length 984;
  Best Local Similarity   48.8%;  Pred. No. 4.1e-06;
  Matches  317;  Conservative    0;  Mismatches  314;  Indels   18;  Gaps    4;

Qy          48 ATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAA 107
               | ||||  | |||| |||   |  | ||  |  ||| |          |    |||   |
Db          63 ACTGTTAATGTTAAGTGCTTGTCAAACAAATAAAAAAACAGCAGATTCTGCAACAACAGA 122

Qy         108 AGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAA 167
               |    || | |||   | |||   | ||||  ||    | |||      |  || || ||
Db         123 AACAACAGCTAAAACGGAAGTCACAGTCAAAGACACCAATGGTCAATTAACCGTTCCCAA 182

Qy         168 ACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGA 227
               | | |||||    || |||||| ||||  ||   |   | ||| |  | | |||  |||
Db         183 AAATCCTAAGAAAGTCGTTGTTTTTGATAATGGTTCCTTGGATACAATGGATGCACTAGG 242

Qy         228 TGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAG 287
               |||          | | |||  |     |  ||||| |  | |     |     |  |
Db         243 TGTCGGTGACCGCGTGGTAGGTGCGCCAACTAAAAATATCCCTGCGTATTTGAAAAAATA 302

Qy         288 AGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAAT 347
                 | ||| ||| |         || | |||      |||  ||||  | |||||| ||||
Db         303 CCAAAAAGTTGAA---------TCAGCAGGCGGCATTAAAGAACCAGATTTAGAAAAAAT 353

Qy         348 CAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAA 407
               || | |||| |||||||| |||||||||  |  |  |  |  ||   ||| |  | |  |
Db         354 CAATCAACTAAAACCAGACTTAATTATTATTTCTGGTCGTCAAC---AAGATTATCAAGA 410

Qy         408 AGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGA 467
               | | |||||    ||||| || || ||| |  |          |    | ||  |   |
Db         411 ACAATTAAAAGCCATTGCGCCAACCATTTACTTAGCTGTAGATGCCAAAAATCCTTGGGC 470

Qy         468 ---AAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAA 524
                  ||  |   |   | |  ||||  |    |    ||| | ||||||||| | || |||
Db         471 ATCAACGAAACAAAATATCGAAACGTTAGGCACTATTTTTGATAAAGAAGAGGTAGCTAA 530

Qy         525 AAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGA 584
               | ||    ||   |     || ||  ||||||  ||    |||||| |   | |    |
Db         531 AGAAAAAATAACTGGCTTAGAAAAAGAAATTGCTGACGTGAAAAAACA---AGCAGAAGC 587

Qy         585 TAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAG 644
               ||      | || |   |   ||   ||| |    ||  |    |    ||| |   ||
Db         588 TAGCGCGAATAATGCGCTTGTTGTGTTAGTTAACGAAGGACAACTTTCCGCTTACGGAAA 647

Qy         645 CAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCA 693
                  ||||  | | |||  | | |    |  | | || || |||| ||||
Db         648 AGGCTCTCGTTTCGGTTTAATTCATGATACATTTGGCTTCAAAGCAGCA 696


RESULT 19
US-09-902-540-1583
; Sequence 1583, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1583
;   LENGTH: 709
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-1583

  Query Match              6.9%;  Score 69.8;  DB 4;  Length 709;
  Best Local Similarity   49.9%;  Pred. No. 5.7e-06;
  Matches  210;  Conservative    0;  Mismatches  202;  Indels    9;  Gaps    1;

Qy     264 AAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACG 323
           |||||      || | |   | | || | | ||   |      |    ||| || |||
Db       7 AAATCCGCGGTTTGACCATATTAAAGGTCAGATGAAAGGCGTTAAAAGTGTGGGTACAGA 66

Qy     324 TAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAA 383
            |   | || | |||||||| ||||  |    | || || || || ||||| |   ||||
Db      67 AAGCGAGCCAAGCTTAGAAGCAATCGCTGCTTTAAAGCCTGACTTGATTATCGGAAATAA 126

Qy     384 TAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAA 443
            |   ||||  ||    | ||  |  |  | | | |||||||||| ||  |        |
Db     127 AATGCGACAAGAAAAAGTATACGATCAGCTTAGTCAAATTGCTCCGACTGTATTCTCTGA 186

Qy     444 AAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTT 503
                |    |||||||  ||   |||          |||||   |  | |||||||| |
Db     187 GGAGCTGCGCGGAGATTGGAAATCAAAC---------TTTAAGCTATATGCAAAAGCTGT 237

Qy     504 AGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATA 563
           |   ||||||||| |||||||| ||   ||||  ||  | || ||     | | ||| |
Db     238 AAACAAAGAAGAAAAAGGTAAAGAAGTATTAGCTGACTATGACAATCGCGTAGCAGATTT 297

Qy     564 TAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATC 623
            || |||    |        |   ||| |||||  | |    || |     | |
Db     298 AAAGAAACGCCTTGGAGATCAGTTAAAACAAAAAATCTCAGTTGTACGCTTTACAGCTGG 357

Qy     624 AGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTT 683
           ||   | | |    |||  |   | || |     |||    |  ||  ||| || || ||
Db     358 AGACGTTCGTATCTATCATAAAGATTCATTCTCAGGTGTTATTTTAGATCAGCTTGGATT 417

Qy     684 T 684
           |
Db     418 T 418


RESULT 20
US-09-902-540-2208
; Sequence 2208, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 2208
;   LENGTH: 702
```

```
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-2208

  Query Match              6.8%;  Score 68.8;  DB 4;  Length 702;
  Best Local Similarity   49.8%;  Pred. No. 9.4e-06;
  Matches  209;  Conservative    0;  Mismatches  202;  Indels    9;  Gaps    1;

Qy     265 AATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGT 324
           ||||      || | |   | | || | | ||   |      |    ||| || |||
Db       1 AATCCGCGGTTTGACCATATTAAAGGTCAGATGAAAGGCGTTAAAAGTGTGGGTACAGAA 60

Qy     325 AAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAAT 384
           |   | || | |||||||| ||||  |    | || || || || ||||| |   ||||
Db      61 AGCGAGCCAAGCTTAGAAGCAATCGCTGCTTTAAAGCCTGACTTGATTATCGGAAATAAA 120

Qy     385 AATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAA 444
           |   ||||  ||    | ||  |  |  | | | |||||||||| ||  |        |
Db     121 ATGCGACAAGAAAAAGTATACGATCAGCTTAGTCAAATTGCTCCGACTGTATTCTCTGAG 180

Qy     445 AGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTA 504
               |    |||||||  ||   |||          |||||   |  | |||||||| ||
Db     181 GAGCTGCGCGGAGATTGGAAATCAAAC---------TTTAAGCTATATGCAAAAGCTGTA 231

Qy     505 GGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATAT 564
              ||||||||| |||||||| ||   ||||  ||  | || ||     | | ||| |
Db     232 AACAAAGAAGAAAAAGGTAAAGAAGTATTAGCTGACTATGACAATCGCGTAGCAGATTTA 291

Qy     565 AAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCA 624
           || |||    |        |   ||| |||||  | |    || |     | |      |
Db     292 AAGAAACGCCTTGGAGATCAGTTAAAACAAAAAATCTCAGTTGTACGCTTTACAGCTGGA 351

Qy     625 GGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTT 684
           |   | | |    |||  |   | || |     |||    |  ||  ||| || || |||
Db     352 GACGTTCGTATCTATCATAAAGATTCATTCTCAGGTGTTATTTTAGATCAGCTTGGATTT 411


RESULT 21
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-Fls
US-08-232-463-14

  Query Match              6.8%;  Score 68.8;  DB 1;  Length 7218;
  Best Local Similarity   4.0%;  Pred. No. 1.5e-05;
  Matches   16;  Conservative  235;  Mismatches  147;  Indels    0;  Gaps    0;

Qy        201 ATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAA 260
              | | ||| | | | |||  |  || | |   :::   :  :::: :: ::: :: :: ::
Db       1464 AGTAGTTAAAGAGATAGAAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1405

Qy        261 AAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAAC 320
              ::::::  : :  :  :::  :  :::::: ::::  :::::: : :     : :::::
Db       1404 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1345

Qy        321 ACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGA 380
              : : ::: ::   ::   ::::::::  :: :::   :::  :::   ::  :  :  ::
Db       1344 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1285

Qy        381 TAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACT 440
               :: :: ::: : ::::: :   : :::::   ::: ::::  :  :     :  :: :::
Db       1284 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1225

Qy        441 GAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGC 500
              ::::::    :: :::::  : :: ::::: :  :: :     :::: ::    :::::
Db       1224 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1165

Qy        501 TTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGA 560
                 ::: ::::::::::::::: :::::: :   ::::::: : :: :::::::  :::::
Db       1164 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1105

Qy        561 ATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGG 598
              : : :::::::::: ::  : ::: :::::  ::::::
Db       1104 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1067

RESULT 22
US-09-071-035-39
; Sequence 39, Application US/09071035
; Patent No. 6448043
;  GENERAL INFORMATION:
;    APPLICANT:  Gil H. Choi
;    TITLE OF INVENTION:  Enterococcus faecalis Polynucleotides and Polypeptides
;    NUMBER OF SEQUENCES:  496
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/071,035
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  A. Anders Brookes
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB369P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  889 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-09-071-035-39

  Query Match              6.5%;  Score 65.6;  DB 3;  Length 889;
  Best Local Similarity   49.3%;  Pred. No. 4.8e-05;
  Matches  294;  Conservative    0;  Mismatches  284;  Indels   18;  Gaps    4;

Qy        101 CAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAG 160
              |||   ||    || | |||   | |||   | ||||  ||    | |||      |  |
Db         36 CAACAGAAACAACAGCTAAAACGGAAGTCACAGTCAAAGACACCAATGGTCAATTAACCG 95

Qy        161 TACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTG 220
              | || ||| | |||||    || |||||| ||||  ||   |   | ||| |  | | ||
Db         96 TTCCCAAAAATCCTAAGAAAGTCGTTGTTTTTGATAATGGTTCCTTGGATACAATGGATG 155

Qy        221 CTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAAC 280
              |  ||| |||          | | |||  |     |  ||||| |  | |     |
Db        156 CACTAGGTGTCGGTGACCGCGTGGTAGGTGCGCCAACTAAAAATATCCCTGCGTATTTGA 215

Qy        281 CATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAG 340
               |  |    | ||| ||| |         || | |||      |||  ||||  | ||||
Db        216 AAAAATACCAAAAAGTTGAA---------TCAGCAGGCGGCATTAAAGAACCAGATTTAG 266

Qy        341 AAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTA 400
              || |||||| | |||| |||||||| ||||||||   |  |  |  | | ||  ||| |
Db        267 AAAAAATCAATCAACTAAAACCAGACTTAATTAT---TATTTCTGGTCGTCAACAAGATT 323

Qy        401 TTTATAAAGACTTAAATAAAATTGCTCCTACGATTGA---ACTGAAAAGTTTCGATGGAG 457
               | |  || | |||||    ||||| || || ||| |   |     |  |  | |
Db        324 ATCAAGAACAATTAAAAGCCATTGCGCCAACCATTTACTTAGCTGTAGATGCCAAAAATC 383

Qy        458 ATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAG 517
               ||       ||  |   |   | |  ||||  |    |       ||| | |||||||| |
Db        384 CTTGGGCATCAACGAAACAAAATATCGAAACGTTAGGCACTATTTTTGATAAAGAAGAGG 443

Qy        518 AAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAA 577
               || |||| ||    ||   |     || ||  ||||||  ||    |||||| |   |
Db        444 TAGCTAAAGAAAAAATAACTGGCTTAGAAAAAGAAATTGCTGACGTGAAAAAACA---AG 500

Qy        578 CTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTC 637
              |    | ||      | || |   |   ||   ||| |    ||  |    |    |||
Db        501 CAGAAGCTAGCGCGAATAATGCGCTTGTTGTGTTAGTTAACGAAGGACAACTTTCCGCTT 560

Qy        638 ATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCA 693
              |   ||    ||||  | | |||  | | |    |  | | || || |||| ||||
Db        561 ACGGAAAAGGCTCTCGTTTCGGTTTAATTCATGATACATTTGGCTTCAAAGCAGCA 616

RESULT 23
US-09-710-279-3348/c
; Sequence 3348, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
;  APPLICANT: KIMMERLY, WILLIAM JOHN
;  TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;  FILE REFERENCE: PU3480US
;  CURRENT APPLICATION NUMBER: US/09/710,279
;  CURRENT FILING DATE:  2000-11-09
;  PRIOR APPLICATION NUMBER: 60/164,258
;  PRIOR FILING DATE: 1999-11-09
;  NUMBER OF SEQ ID NOS: 4472
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3348
;   LENGTH: 3153
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3348

  Query Match             6.4%;  Score 64.6;  DB 4;  Length 3153;
  Best Local Similarity   47.0%;  Pred. No. 0.0001;
  Matches  452;  Conservative    0;  Mismatches  479;  Indels   30;  Gaps    7;

Qy          61 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAA 120
               | ||   ||  || |  || ||  |  |     |||     |   |  |    |  |||
Db        1188 AGTGATTCTAAAGAAACTGTAACCATCAAAAATAGTTTTGAAGCAAGTGGTAAAGAAAAT 1129

Qy         121 GATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGT 180
                ||||   ||| |  ||  |         ||||    ||||||| ||  | ||||||  |
Db        1128 AATGGCAGTGATAAGAAAAAAATCTCTAATACTGTCGAAGTACCAAAGAATCCTAAAAAT 1069

Qy         181 GTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGT------TAAA 234
               |  |||||  | || |||  |    |||||| |||    |  |||| |||      ||||
Db        1068 GCCGTTGTATTAGATTATGGAGCGCTTGATGTGTTGAAAGAATTAGGTGTGGCTGATAAA 1009

Qy         235 CCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGA---T 291
                     ||  || |  | | | | || ||  |   | |   ||||   |||   ||   |
Db        1008 GTAAAAGGTTTACCTAAAGGTGAAAATAACCAATCTTTACCTAAATTTTTAGATGAATTT 949

Qy         292 AAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGT 351
               |||  ||  || || | |  |   ||||     ||  ||   ||||| ||  || |
Db         948 AAAGATGATAAGTATATTAATACTGGAAATTTAAAAGAAGTGAACTTTGATAAAGTTGCA 889

Qy         352 AAACTTAAACCAGATTTAA---TTATTGCTGATAATAATAGACACAAAGGTATTTATAAA 408
                 |  |||||||||| | |   ||||| | |  |  |      |  |    | ||   |
Db         888 TCAGCTAAACCAGATGTGATTTTTATTTCAGGAAGAACAGCTAATCAGAAAAATTTAGAT 829

Qy         409 GACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAA 468
               || || || |||  ||| || |    | || |   |  | | | |  |  | || |  |
Db         828 GAATTTAAAAAAGCTGCACCAAAAGCTAAAGTTGTATATGTAGGTACAAGTGATGACAAC 769

Qy         469 AATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAA 528
                  ||| | | | | |||| || | || ||  |||||| |||    | ||   |||| ||
Db         768 TTAATTAAAGATATGAAAAAAAATACAGAAAATTTAGGGAAAATCTACGA---TAAAGAA 712

Qy         529 CGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAA 588
                    ||   || |   ||| |||  |  |||    |||| |      |  |  ||||||
Db         711 GATAAAGCTAAAAAAATTAATAAAGATTTAGATAGAAAAATATCTGATATGAAAGATAAA 652

Qy         589 AATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAAC 648
               | | || |  |         ||||     |  ||  | |      | | | | |   |
Db         651 ACTAAAGACTTTAATAAGAAAGTAATGTATTTATTGGTTAACGAAGGTGAACTATCAACG 592

Qy         649 TCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTT 708
               | |      ||   |    |  ||    |||  ||| |   |  |  |  | |   ||
Db         591 TTTGGACCAGGAGGAAGATTTGGTGGTTTAGTGTTTGATACATTAGGATTTAAACCTGCA 532

Qy         709 ACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTA 768
                  |||    | || ||    |   | || |       | | || || | ||||  | |
Db         531 GACAAAAAGGTTAGCAAAAGCCCGCATGGTC------AAAATATAAATAATGAATATATT 478

Qy         769 TCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACG---AA 825
                   |   |||||| ||   ||| ||    |||          ||  |      |   ||
Db         477 AACAAGCAGAATCCAGATGTTATTTTAGCTATGGATCGTGGTTCAGTTGTAGGTGGTAAA 418

Qy         826 CCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAAT 885
                |  ||  |||  || |   |||| |    ||    || ||  |||| || || |||| |
Db         417 GCAACAACAAATCAAGTTTTAAAAAACAAAGTTATAAAAAATGTAAAAGCAGTAAAAAGT 358

Qy         886 CAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCA 945
                | | | || |    ||||| |   |  |||||      ||      || |   ||||||
Db         357 AATCATATTTACGAATTAGATCCAAAACTATGGTATTTCTC------TTCAGGATCTTCA 304

Qy         946 GAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATAAG 1005
                    ||     ||| || ||| |||||||  | |   ||     |||   || ||  |
Db         303 ACGACAACTATCAAACAAATTGATGAATTAAATGAAGTAGTAGAGAAAGTTGAAAAATAA 244

Qy        1006 T 1006
               |
Db         243 T 243

RESULT 24
US-09-710-279-4042
; Sequence 4042, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
;  APPLICANT: KIMMERLY, WILLIAM JOHN
;  TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;  FILE REFERENCE: PU3480US
;  CURRENT APPLICATION NUMBER: US/09/710,279
;  CURRENT FILING DATE:  2000-11-09
;  PRIOR APPLICATION NUMBER: 60/164,258
;  PRIOR FILING DATE: 1999-11-09
;  NUMBER OF SEQ ID NOS: 4472
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4042
;   LENGTH: 3267
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: nucleic acid sequence
US-09-710-279-4042

  Query Match             6.4%;  Score 64.6;  DB 4;  Length 3267;
  Best Local Similarity   47.0%;  Pred. No. 0.0001;
  Matches  452;  Conservative    0;  Mismatches  479;  Indels   30;  Gaps    7;

Qy          61 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAA 120
               | ||   ||  || |  || ||  |  |     |||     |   |  |    |  |||
Db         800 AGTGATTCTAAAGAAACTGTAACCATCAAAAATAGTTTTGAAGCAAGTGGTAAAGAAAAT 859

Qy         121 GATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGT 180
                ||||   ||| |  ||  |         ||||    ||||||| ||  | ||||||  |
Db         860 AATGGCAGTGATAAGAAAAAAATCTCTAATACTGTCGAAGTACCAAAGAATCCTAAAAAT 919

Qy         181 GTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGT------TAAA 234
               |  |||||  | || |||  |    |||||| |||    |  |||| |||      ||||
Db         920 GCCGTTGTATTAGATTATGGAGCGCTTGATGTGTTGAAAGAATTAGGTGTGGCTGATAAA 979

Qy         235 CCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGA---T 291
                     ||  || |  | | | | || ||  |   | |   ||||   |||   ||   |
Db         980 GTAAAAGGTTTACCTAAAGGTGAAAATAACCAATCTTTACCTAAATTTTTAGATGAATTT 1039

Qy         292 AAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGT 351
               |||  ||  || || | |  |   ||||     ||  ||   ||||| ||  || |
Db        1040 AAAGATGATAAGTATATTAATACTGGAAATTTAAAAGAAGTGAACTTTGATAAAGTTGCA 1099

Qy         352 AAACTTAAACCAGATTTAA---TTATTGCTGATAATAATAGACACAAAGGTATTTATAAA 408
                 |  |||||||||| | |   ||||| | |  |  |      |  |    | ||   |
Db        1100 TCAGCTAAACCAGATGTGATTTTTATTTCAGGAAGAACAGCTAATCAGAAAAATTTAGAT 1159

Qy         409 GACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAA 468
               || || || |||  ||| || |    | || |   |  | | | |  |  | || |  |
Db        1160 GAATTTAAAAAAGCTGCACCAAAAGCTAAAGTTGTATATGTAGGTACAAGTGATGACAAC 1219

Qy         469 AATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAA 528
                  ||| | | | | |||| || | || ||  |||||| |||    | ||   |||| ||
Db        1220 TTAATTAAAGATATGAAAAAAAATACAGAAAATTTAGGGAAAATCTACGA---TAAAGAA 1276

Qy         529 CGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAA 588
                    ||   || |   ||| |||  |  |||    |||| |      |  |  ||||||
Db        1277 GATAAAGCTAAAAAAATTAATAAAGATTTAGATAGAAAAATATCTGATATGAAAGATAAA 1336
```

```
Qy      589 AATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAAC 648
            | | || |  |         ||||     |  ||  | |      | | | | |   |
Db     1337 ACTAAAGACTTTAATAAGAAAGTAATGTATTTATTGGTTAACGAAGGTGAACTATCAACG 1396

Qy      649 TCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTT 708
            | |      ||   |    |  ||    |||  ||| |   |  |  |  | |   ||
Db     1397 TTTGGACCAGGAGGAAGATTTGGTGGTTTAGTGTTTGATACATTAGGATTTAAACCTGCA 1456

Qy      709 ACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTA 768
               |||    | || ||    |   | || |       | | || || | ||||  | |
Db     1457 GACAAAAAGGTTAGCAAAAGCCCGCATGGTC------AAAATATAAATAATGAATATATT 1510

Qy      769 TCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACG---AA 825
                |   |||||| ||   ||| ||    |||          ||  |      |   ||
Db     1511 AACAAGCAGAATCCAGATGTTATTTTAGCTATGGATCGTGGTTCAGTTGTAGGTGGTAAA 1570

Qy      826 CCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAAT 885
             |  ||  |||  || |   |||| |    ||    || ||  |||| || || |||| |
Db     1571 GCAACAACAAATCAAGTTTTAAAAAACAAAGTTATAAAAAATGTAAAAGCAGTAAAAAGT 1630

Qy      886 CAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCA 945
             | | | || |    ||||| |   |  |||||      ||      || |   ||||||
Db     1631 AATCATATTTACGAATTAGATCCAAAACTATGGTATTTCTC------TTCAGGATCTTCA 1684

Qy      946 GAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATAAG 1005
                 ||     ||| || ||| |||||||  | |   ||     |||   || ||  |
Db     1685 ACGACAACTATCAAACAAATTGATGAATTAAATGAAGTAGTAGAGAAAGTTGAAAAATAA 1744

Qy     1006 T 1006
            |
Db     1745 T 1745
```

```
RESULT 25
US-09-710-279-3564
; Sequence 3564, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
;  APPLICANT: KIMMERLY, WILLIAM JOHN
;  TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;  FILE REFERENCE: PU3480US
;  CURRENT APPLICATION NUMBER: US/09/710,279
;  CURRENT FILING DATE:  2000-11-09
;  PRIOR APPLICATION NUMBER: 60/164,258
;  PRIOR FILING DATE: 1999-11-09
;  NUMBER OF SEQ ID NOS: 4472
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3564
;   LENGTH: 3618
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3564

  Query Match             6.4%;  Score 64.6;  DB 4;  Length 3618;
  Best Local Similarity   47.0%;  Pred. No. 0.00011;
  Matches  452;  Conservative    0;  Mismatches  479;  Indels   30;  Gaps    7;

Qy       61 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAA 120
            | ||   ||  || |  || ||  |  |     |||     |   |  |    |  |||
Db      666 AGTGATTCTAAAGAAACTGTAACCATCAAAAATAGTTTTGAAGCAAGTGGTAAAGAAAAT 725

Qy      121 GATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGT 180
             ||||   ||| |  ||  |         ||||    ||||||| ||  | ||||||  |
Db      726 AATGGCAGTGATAAGAAAAAAATCTCTAATACTGTCGAAGTACCAAAGAATCCTAAAAAT 785

Qy      181 GTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGT------TAAA 234
            |  |||||  | || |||  |    |||||| |||    |  |||| |||      ||||
Db      786 GCCGTTGTATTAGATTATGGAGCGCTTGATGTGTTGAAAGAATTAGGTGTGGCTGATAAA 845

Qy      235 CCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGA---T 291
                  ||  || |  | | | | || ||  |   | |   ||||   |||   ||   |
Db      846 GTAAAAGGTTTACCTAAAGGTGAAAATAACCAATCTTTACCTAAATTTTTAGATGAATTT 905

Qy      292 AAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGT 351
            |||  ||  || || | |  |   ||||     ||  ||   ||||| ||  || |
Db      906 AAAGATGATAAGTATATTAATACTGGAAATTTAAAAGAAGTGAACTTTGATAAAGTTGCA 965

Qy      352 AAACTTAAACCAGATTTAA---TTATTGCTGATAATAATAGACACAAAGGTATTTATAAA 408
              |  |||||||||| | |   ||||| | |  |  |      |  |    | ||   |
Db      966 TCAGCTAAACCAGATGTGATTTTTATTTCAGGAAGAACAGCTAATCAGAAAAATTTAGAT 1025

Qy      409 GACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAA 468
            || || || |||  ||| || |    | || |   |  | | | |  |  | || |  |
Db     1026 GAATTTAAAAAAGCTGCACCAAAAGCTAAAGTTGTATATGTAGGTACAAGTGATGACAAC 1085

Qy      469 AATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAA 528
               ||| | | | | |||| || | || ||  |||||| |||    | ||   |||| ||
Db     1086 TTAATTAAAGATATGAAAAAAAATACAGAAAATTTAGGGAAAATCTACGA---TAAAGAA 1142

Qy      529 CGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAA 588
                 ||   || |   ||| |||  |  |||    |||| |      |  |  ||||||
Db     1143 GATAAAGCTAAAAAAATTAATAAAGATTTAGATAGAAAAATATCTGATATGAAAGATAAA 1202

Qy      589 AATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAAC 648
            | | || |  |         ||||     |  ||  | |      | | | | |   |
Db     1203 ACTAAAGACTTTAATAAGAAAGTAATGTATTTATTGGTTAACGAAGGTGAACTATCAACG 1262

Qy      649 TCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTT 708
            | |      ||   |    |  ||    |||  ||| |   |  |  |  | |   ||
Db     1263 TTTGGACCAGGAGGAAGATTTGGTGGTTTAGTGTTTGATACATTAGGATTTAAACCTGCA 1322

Qy      709 ACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTA 768
               |||    | || ||    |   | || |       | | || || | ||||  | |
Db     1323 GACAAAAAGGTTAGCAAAAGCCCGCATGGTC------AAAATATAAATAATGAATATATT 1376

Qy      769 TCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACG---AA 825
                |   |||||| ||   ||| ||    |||          ||  |      |   ||
Db     1377 AACAAGCAGAATCCAGATGTTATTTTAGCTATGGATCGTGGTTCAGTTGTAGGTGGTAAA 1436

Qy      826 CCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAAT 885
             |  ||  |||  || |   |||| |    ||    || ||  |||| || || |||| |
Db     1437 GCAACAACAAATCAAGTTTTAAAAAACAAAGTTATAAAAAATGTAAAAGCAGTAAAAAGT 1496

Qy      886 CAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCA 945
             | | | || |    ||||| |   |  |||||      ||      || |   ||||||
Db     1497 AATCATATTTACGAATTAGATCCAAAACTATGGTATTTCTC------TTCAGGATCTTCA 1550

Qy      946 GAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATAAG 1005
                 ||     ||| || ||| |||||||  | |   ||     |||   || ||  |
Db     1551 ACGACAACTATCAAACAAATTGATGAATTAAATGAAGTAGTAGAGAAAGTTGAAAAATAA 1610

Qy     1006 T 1006
            |
Db     1611 T 1611
```

```
RESULT 26
US-09-710-279-461
; Sequence 461, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
;  APPLICANT: KIMMERLY, WILLIAM JOHN
;  TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;  FILE REFERENCE: PU3480US
;  CURRENT APPLICATION NUMBER: US/09/710,279
;  CURRENT FILING DATE:  2000-11-09
;  PRIOR APPLICATION NUMBER: 60/164,258
;  PRIOR FILING DATE: 1999-11-09
;  NUMBER OF SEQ ID NOS: 4472
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 461
;   LENGTH: 1044
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: nucleic acid sequence
US-09-710-279-461

  Query Match             6.4%;  Score 64.4;  DB 4;  Length 1044;
  Best Local Similarity   47.1%;  Pred. No. 9e-05;
  Matches  450;  Conservative    0;  Mismatches  476;  Indels   30;  Gaps    7;

Qy         61 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAA 120
              | ||   ||  ||  |  ||  ||  |  |     |||     |   |  |    |  |||
Db        100 AGTGATTCTAAAGAAACTGTAACCATCAAAAATAGTTTTGAAGCAAGTGGTAAAGAAAAT 159

Qy        121 GATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGT 180
               ||||   ||| |  ||  |         ||||    ||||||| ||  | ||||||  |
Db        160 AATGGCAGTGATAAGAAAAAAATCTCTAATACTGTCGAAGTACCAAAGAATCCTAAAAAT 219

Qy        181 GTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGT------TAAA 234
              |  |||||  | || |||  |    |||||| |||    |  |||| |||      ||||
Db        220 GCCGTTGTATTAGATTATGGAGCGCTTGATGTGTTGAAAGAATTAGGTGTGGCTGATAAA 279

Qy        235 CCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGA---T 291
                    ||  || |  | | | | || ||  |   | |   ||||   |||   ||   |
Db        280 GTAAAAGGTTTACCTAAAGGTGAAAATAACCAATCTTTACCTAAATTTTTAGATGAATTT 339

Qy        292 AAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGT 351
              |||  ||  || || | |  |   ||||     ||  ||   ||||| ||  || |
Db        340 AAAGATGATAAGTATATTAATACTGGAAATTTAAAAGAAGTGAACTTTGATAAAGTTGCA 399

Qy        352 AAACTTAAACCAGATTTAA---TTATTGCTGATAATAATAGACACAAAGGTATTTATAAA 408
                |  |||||||||| | |   ||||| | |  |  |      |  |    | ||   |
Db        400 TCAGCTAAACCAGATGTGATTTTTATTTCAGGAAGAACAGCTAATCAGAAAAATTTAGAT 459

Qy        409 GACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAA 468
              || || || |||  ||| || |    | || |   |  | | | |  |  | || |  |
Db        460 GAATTTAAAAAAGCTGCACCAAAAGCTAAAGTTGTATATGTAGGTACAAGTGATGACAAC 519

Qy        469 AATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAA 528
                 ||| | | | | |||| || | || ||  |||||| |||    | ||   |||| ||
Db        520 TTAATTAAAGATATGAAAAAAAATACAGAAAATTTAGGGAAAATCTACGA---TAAAGAA 576

Qy        529 CGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAA 588
                   ||   || |   ||| |||  |  |||    |||| |      |  |  ||||||
Db        577 GATAAAGCTAAAAAAATTAATAAAGATTTAGATAGAAAAATATCTGATATGAAAGATAAA 636

Qy        589 AATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAAC 648
              | | || |  |         ||||     |  ||  | |      | | | | | |   |
Db        637 ACTAAAGACTTTAATAAGAAAGTAATGTATTTATTGGTTAACGAAGGTGAACTATCAACG 696

Qy        649 TCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTT 708
              | |      ||   |    |  ||    |||  ||| |   |  |  |  | |   ||
Db        697 TTTGGACCAGGAGGAAGATTTGGTGGTTTAGTGTTTGATACATTAGGATTTAAACCTGCA 756

Qy        709 ACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTA 768
                 |||    | || ||    |   | || |       | | || || | ||||  | |
Db        757 GACAAAAAGGTTAGCAAAAGCCCGCATGGTC------AAAATATAAATAATGAATATATT 810

Qy        769 TCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACG---AA 825
                  |   |||||| ||   ||| ||    |||          ||  |      |   ||
Db        811 AACAAGCAGAATCCAGATGTTATTTTAGCTATGGATCGTGGTTCAGTTGTAGGTGGTAAA 870

Qy        826 CCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAAT 885
               |  ||  |||  || |   |||| |    ||    || ||  |||| || || |||| |
Db        871 GCAACAACAAATCAAGTTTTAAAAAACAAAGTTATAAAAAATGTAAAAGCAGTAAAAAGT 930

Qy        886 CAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCA 945
               | | | || |    ||||| |   |  |||||      ||      || |   ||||||
Db        931 AATCATATTTACGAATTAGATCCAAAACTATGGTATTTCTC------TTCAGGATCTTCA 984

Qy        946 GAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAA 1001
                   ||     ||| || ||| |||||||  | |   ||     |||   || ||
Db        985 ACGACAACTATCAAACAAATTGATGAATTAAATGAAGTAGTAGAGAAAGTTGAAAA 1040
```

---

```
RESULT 27
US-09-710-279-1267
; Sequence 1267, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
;  APPLICANT: KIMMERLY, WILLIAM JOHN
;  TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;  FILE REFERENCE: PU3480US
;  CURRENT APPLICATION NUMBER: US/09/710,279
;  CURRENT FILING DATE:  2000-11-09
;  PRIOR APPLICATION NUMBER: 60/164,258
;  PRIOR FILING DATE: 1999-11-09
;  NUMBER OF SEQ ID NOS: 4472
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1267
;   LENGTH: 1044
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: nucleic acid sequence
US-09-710-279-1267

  Query Match             6.4%;  Score 64.4;  DB 4;  Length 1044;
  Best Local Similarity   47.1%;  Pred. No. 9e-05;
  Matches  450;  Conservative    0;  Mismatches  476;  Indels   30;  Gaps    7;

Qy         61 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAA 120
              | ||   ||  ||  |  ||  ||  |  |     |||     |   |  |    |  |||
Db        100 AGTGATTCTAAAGAAACTGTAACCATCAAAAATAGTTTTGAAGCAAGTGGTAAAGAAAAT 159

Qy        121 GATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGT 180
               ||||   ||| |  ||  |         ||||    ||||||| ||  | ||||||  |
Db        160 AATGGCAGTGATAAGAAAAAAATCTCTAATACTGTCGAAGTACCAAAGAATCCTAAAAAT 219

Qy        181 GTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGT------TAAA 234
              |  |||||  | || |||  |    |||||| |||    |  |||| |||      ||||
Db        220 GCCGTTGTATTAGATTATGGAGCGCTTGATGTGTTGAAAGAATTAGGTGTGGCTGATAAA 279

Qy        235 CCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGA---T 291
                    ||  || |  | | | | || ||  |   | |   ||||   |||   ||   |
Db        280 GTAAAAGGTTTACCTAAAGGTGAAAATAACCAATCTTTACCTAAATTTTTAGATGAATTT 339

Qy        292 AAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGT 351
              |||  ||  || || | |  |   ||||     ||  ||   ||||| ||  || |
Db        340 AAAGATGATAAGTATATTAATACTGGAAATTTAAAAGAAGTGAACTTTGATAAAGTTGCA 399

Qy        352 AAACTTAAACCAGATTTAA---TTATTGCTGATAATAATAGACACAAAGGTATTTATAAA 408
                |  |||||||||| | |   ||||| | |  |  |      |  |    | ||   |
Db        400 TCAGCTAAACCAGATGTGATTTTTATTTCAGGAAGAACAGCTAATCAGAAAAATTTAGAT 459

Qy        409 GACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAA 468
              || || || |||  ||| || |    | || |   |  | | | |  |  | || |  |
Db        460 GAATTTAAAAAAGCTGCACCAAAAGCTAAAGTTGTATATGTAGGTACAAGTGATGACAAC 519

Qy        469 AATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAA 528
                 ||| | | | | |||| || | || ||  |||||| |||    | ||   |||| ||
Db        520 TTAATTAAAGATATGAAAAAAAATACAGAAAATTTAGGGAAAATCTACGA---TAAAGAA 576

Qy        529 CGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAA 588
                   ||   || |   ||| |||  |  |||    |||| |      |  |  ||||||
Db        577 GATAAAGCTAAAAAAATTAATAAAGATTTAGATAGAAAAATATCTGATATGAAAGATAAA 636

Qy        589 AATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAAC 648
              | | || |  |         ||||     |  ||  | |      | | | | | |   |
Db        637 ACTAAAGACTTTAATAAGAAAGTAATGTATTTATTGGTTAACGAAGGTGAACTATCAACG 696

Qy        649 TCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTT 708
              | |      ||   |    |  ||    |||  ||| |   |  |  |  | |   ||
Db        697 TTTGGACCAGGAGGAAGATTTGGTGGTTTAGTGTTTGATACATTAGGATTTAAACCTGCA 756
```

```
Qy        709 ACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTA 768
                 |||    | || ||    |   | || |       | | || || | ||||  | |
Db        757 GACAAAAAGGTTAGCAAAAGCCCGCATGGTC------AAAATATAAATAATGAATATATT 810

Qy        769 TCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACG---AA 825
                  |   |||||| ||   ||| ||    |||          ||  |      |   ||
Db        811 AACAAGCAGAATCCAGATGTTATTTTAGCTATGGATCGTGGTTCAGTTGTAGGTGGTAAA 870

Qy        826 CCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAAT 885
               |  ||  |||  || |   |||| |    ||    || ||  |||| || || |||| |
Db        871 GCAACAACAAATCAAGTTTTAAAAAACAAAGTTATAAAAAATGTAAAAGCAGTAAAAAGT 930

Qy        886 CAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCA 945
               | | | || |    ||||| |   |  |||||      ||      || |   ||||||
Db        931 AATCATATTTACGAATTAGATCCAAAACTATGGTATTTCTC------TTCAGGATCTTCA 984

Qy        946 GAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAA 1001
                   ||     ||| || ||| |||||||  | |   ||     |||   || ||
Db        985 ACGACAACTATCAAACAAATTGATGAATTAAATGAAGTAGTAGAGAAAGTTGAAAA 1040


RESULT 28
US-09-134-001C-2655
; Sequence 2655, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2655
;   LENGTH: 1056
;   TYPE: DNA
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2655

  Query Match             6.4%;  Score 64.4;  DB 3;  Length 1056;
  Best Local Similarity   47.1%;  Pred. No. 9e-05;
  Matches  450;  Conservative    0;  Mismatches  476;  Indels   30;  Gaps    7;

Qy         61 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAA 120
              | ||   ||  || |  || ||  |  |     |||     |   |  |    |  |||
Db        112 AGTGATTCTAAAGAAACTGTAACCATCAAAAATAGTTTTGAAGCAAGTGGTAAAGAAAAT 171

Qy        121 GATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGT 180
               ||||   ||| |  ||  |         ||||    ||||||| ||  | ||||||  |
Db        172 AATGGCAGTGATAAGAAAAAAATCTCTAATACTGTCGAAGTACCAAAGAATCCTAAAAAT 231

Qy        181 GTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGT------TAAA 234
              |  |||||  | || |||  |    ||||||| |||    |  |||| |||      ||||
Db        232 GCCGTTGTATTAGATTATGGAGCGCTTGATGTGTTGAAAGAATTAGGTGTGGCTGATAAA 291

Qy        235 CCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGA---T 291
                    ||  || |  | | | | || ||  |   | |   ||||   |||   ||   |
Db        292 GTAAAAGGTTTACCTAAAGGTGAAAATAACCAATCTTTACCTAAATTTTTAGATGAATTT 351

Qy        292 AAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGT 351
              |||  ||  || || | |  |   ||||     ||  ||   ||||| ||  || |
Db        352 AAAGATGATAAGTATATTAATACTGGAAATTTAAAAGAAGTGAACTTTGATAAAGTTGCA 411

Qy        352 AAACTTAAACCAGATTTAA---TTATTGCTGATAATAATAGACACAAAGGTATTTATAAA 408
                |  |||||||||| | |   ||||| | |  |  |      |  |    | ||   |
Db        412 TCAGCTAAACCAGATGTGATTTTTATTTCAGGAAGAACAGCTAATCAGAAAAATTTAGAT 471

Qy        409 GACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAA 468
              || || || |||  ||| || |    | || |   |  | | | |  |  | || |  |
Db        472 GAATTTAAAAAAGCTGCACCAAAAGCTAAAGTTGTATATGTAGGTACAAGTGATGACAAC 531

Qy        469 AATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAA 528
                 ||| | | | | |||| || | || ||  |||||| |||    | ||   |||| ||
Db        532 TTAATTAAAGATATGAAAAAAAATACAGAAAATTTAGGGAAAATCTACGA---TAAAGAA 588

Qy        529 CGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAA 588
                   ||   || |   ||| |||  |  |||    |||| |      |  |  ||||||
Db        589 GATAAAGCTAAAAAAATTAATAAAGATTTAGATAGAAAAATATCTGATATGAAAGATAAA 648

Qy        589 AATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAAC 648
              | | || |  |         ||||     |  ||  | |      | | | | |   |
Db        649 ACTAAAGACTTTAATAAGAAAGTAATGTATTTATTGGTTAACGAAGGTGAACTATCAACG 708

Qy        649 TCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTT 708
              | |      ||   |    |  ||    |||  ||| |   |  |  |  | |   ||
Db        709 TTTGGACCAGGAGGAAGATTTGGTGGTTTAGTGTTTGATACATTAGGATTTAAACCTGCA 768

Qy        709 ACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTA 768
                 |||    | || ||    |   | || |       | | || || | ||||  | |
Db        769 GACAAAAAGGTTAGCAAAAGCCCGCATGGTC------AAAATATAAATAATGAATATATT 822

Qy        769 TCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACG---AA 825
                  |   |||||| ||   ||| ||    |||          ||  |      |   ||
Db        823 AACAAGCAGAATCCAGATGTTATTTTAGCTATGGATCGTGGTTCAGTTGTAGGTGGTAAA 882

Qy        826 CCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAAT 885
               |  ||  |||  || |   |||| |    ||    || ||  |||| || || |||| |
Db        883 GCAACAACAAATCAAGTTTTAAAAAACAAAGTTATAAAAAATGTAAAAGCAGTAAAAAGT 942

Qy        886 CAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCA 945
               | | | || |    ||||| |   |  |||||      ||      || |   ||||||
Db        943 AATCATATTTACGAATTAGATCCAAAACTATGGTATTTCTC------TTCAGGATCTTCA 996

Qy        946 GAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAA 1001
                   ||     ||| || ||| |||||||  | |   ||     |||   || ||
Db        997 ACGACAACTATCAAACAAATTGATGAATTAAATGAAGTAGTAGAGAAAGTTGAAAA 1052


RESULT 29
US-09-902-540-1280
; Sequence 1280, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1280
;   LENGTH: 1039
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(1039)
;   OTHER INFORMATION: unsure at all n locations
US-09-902-540-1280

  Query Match             6.3%;  Score 63.6;  DB 4;  Length 1039;
  Best Local Similarity   48.1%;  Pred. No. 0.00013;
  Matches  177;  Conservative    0;  Mismatches  191;  Indels    0;  Gaps    0;

Qy        254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
              | || |||||||||      |  ||| ||  || | | |||||   |||| | |    |
Db        268 AAAAAAAAAAAAATAAATAAAAAAAATCAAAAAAAAAAAAAATAAAAAAAAAAAAAATTA 327
```

```
Qy          314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
                 |   |  | |||   |   ||  |  || |||| |  |||   | |  | ||| || ||
Db          328 AAATCAATCATAAAATAAAAAAAATTAAATAAATAACAAAAAAAACAAAAAAATTAAAATA 387

Qy          374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                       ||||     ||| ||    |   |   | |   ||| ||||       |   |
Db          388 AAAAAATTAATTTACAACAAAATAAAAGAAAANTAAAAAAAAAAAAAAAAATAATATGTAA 447

Qy          434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                   ||    ||       |   | |  | ||  | ||     |       |||  ||
Db          448 AAAAAAAACAACAAAAACACCAAAAAAAAAAAAAATAAATAAAAAAAAAAAAAAATGAAAAC 507

Qy          494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
                 ||||  |  |  |||| ||  | ||   |||||||   || || || |  | || || |
Db          508 AAAAAAATAAAAATAAAAAAATAAAAAAAAAAAAAACAAATAAAAAAAAAAATAAAAAAAGA 567

Qy          554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
                |  || || | |||||  |||  ||      |||||| |||||   |    | |||  |
Db          568 TAAAAAAAAAAAAAAAAGTAAAATACATAATTTAAAAAACAAAAAAAAAACNCCGCATAAA 627

Qy          614 CTGCTAAA 621
                 |   |||
Db          628 ATATAAAA 635

RESULT 30
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
;  APPLICANT: The University of British Columbia
;  TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
;  FILE REFERENCE: 4810-58741
;  CURRENT APPLICATION NUMBER: US/09/806,708B
;  CURRENT FILING DATE:  2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/147,133
;  PRIOR FILING DATE: 1999-08-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
;   LENGTH: 1141
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   NAME/KEY: promoter
;   LOCATION: (1)..(1141)
;   OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

  Query Match              6.2%;  Score 62.2;  DB 4;  Length 1141;
  Best Local Similarity   11.7%;  Pred. No. 0.00027;
  Matches  117;  Conservative  358;  Mismatches  522;  Indels    4;  Gaps    2;

Qy           10 TCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACT 69
                :: :  | ::  |: :::  :::| :::||::   | ::    :  |  :|
Db         1057 WYTRSNANWSCATKBMWWTMKWYATKYRTAWYAMWCAWRNNNMWCATNGYAKSCATNNAM 998

Qy           70 GCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTT 129
                 :|  ::    || :|::||  :::  ||    :| :|| |    : |: ||
Db          997 WYATTRWAAYAAAKWARWAGNNMRMYGAAAGNKWGCMAAMATMGBWWADTAGKMCNNNNN 938

Qy          130 GAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTT 189
                 :  ::::  :            |||  :| |    | : :   : |::: :|  :|:|
Db          937 NWTTDVRRMAMKAKNNNNNNAYWTACYNRAATNNKMATHWMKWTHGAHSKRRTRHHTRTC 878

Qy          190 CTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCG 249
                  |:       | |  :: :   ::   ::  |          |   |:|:  :::::
Db          877 RRTKYNNNNNNARTVYWYHHAARRWMNAWWTRTNNNNNNNNNNACRNTRTWWABWKHSWC 818

Qy          250 GATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACT 309
                               : :: ||   :  :|     ||:: :|     :  :   :  :|:
Db          817 NNNNNNNNNNNNNNNTWCHYTTANABBCYRANNNNAAARMARTCNNYMHAAVTTTHTDWCY 758

Qy          310 TCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTA 369
                : :     ::::        : :|            ::          : |::| ||  :
Db          757 KTWMNTWYWDMMTTMBTTTTRNMTTSTNMTNNNNNNMWACTNNNNNNMWKAYYAHATNNW 698

Qy          370 ATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCT 429
                  ::     ::|  | : ::: :     ::  ||: ::    |
Db          697 GCWWNNTDARRTNNTTVMRRRWMTNTKTRWYSTTRRHHYTGATNNNNNNNNNNNNNNNS 638

Qy          430 ACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACA 489
                 |  |  : :   :  :  |:: ::::::|| | ::::|: :::: |  :   ::    :
Db          637 CCTCTRMMTMRWTMKGDGMTVRKKVKWRDTTCTYVDVWADSWVWWYANWMRCRDV-TYTR 579

Qy          490 ATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAG 549
                  |:|   | ::::::  |:: ::  |: ::: |::  :   :  | :::: |  ::::
Db          578 NNTYCKSYAHSYWYWSNNAMWYRRYSARNWSSMARWTTRNNWWMSGBVRMRWAGTMWWRH 519

Qy          550 AAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCA 609
                :    |: : ::::::|:    : : :| |:: :    :: : :::    : | ::
Db          518 WNNNNTDTRYYWWWKRWARBTTTVYDSMCNAKSMWRGNNWRAMKMWWAANNDAGAMDHWT 459

Qy          610 GTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTA 669
                 ::|   :::::         |   : |     : |:::::: :  |::    |
Db          458 YWMGNNTMWMRRAWKMMNMAWCRRAYCCNNNNNRACVWHKHKMWRWTWKYMWKAACNNNN 399

Qy          670 AGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTAT 729
                 : : :  |   :  :    : :::| : : ::   ::|:: |:       :::   | :
Db          398 BKAMYMRVAWMMYSRDTTNTDWMMWTSDWBWHWYTVDYTMMRAWNNNNNNNNWRBCKTTSW 339

Qy          730 CTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCG- 788
                ::::::   :||:   |:  | : ||: :: | |: :  : :|:: | : : :|
Db          338 MWWMDHMNTHCTYGNNTWGSAYBMAAMSMWAAGASNBVTYNWCWRMTYMGKTMTNNNNNN 279

Qy          789 --TATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAA 846
                  : ::|: : :   ::  : |:: : : :::  |::: :: ::  ::: : :: ::: :
Db          278 KAWYYRTKTVAWCNNRYYYDTAVWTBKRNYKYCYAYBWYYBMYMGKHHWBWWRRABHRSW 219

Qy          847 AAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGAC 906
                 :: ::|  ::: : :  |:::: :|: :     : ::  :::   |    :::::| ::
Db          218 NMWWVKCRNKYMVSWHYHAMRYBKWABAVGCNNNWKDRMAHHHWCATNNNMMWWWYAYMH 159

Qy          907 CGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTT 966
                : ::  : |:     |::: :: :::|:: : ::::  : :: | |::: |: : : :|:
Db          158 HMHKKGKAAWTNNKTABRDDHBAHVKTYWYWRYDYWCAMCWMNAKAKVRTAMKHMWYYTD 99

Qy          967 GTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATAAGTA 1007
                :: :|  |   ::: ::  :: :   ::  :: : : : ::
Db           98 RYVSANNTGVRWMMMRWCMWWYSMNNRWYYRMGRKYTWAWM 58

RESULT 31
US-09-543-681A-1393
; Sequence 1393, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
;  APPLICANT: GARY BRETON
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABI
;  TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 2709.1002-001
;  CURRENT APPLICATION NUMBER: US/09/543,681A
;  CURRENT FILING DATE:  2000-04-05
;  PRIOR APPLICATION NUMBER: US 60/128,706
;  PRIOR FILING DATE: 1999-04-09
;  NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 1393
;   LENGTH: 978
;   TYPE: DNA
;   ORGANISM: Proteus mirabilis
US-09-543-681A-1393

  Query Match              6.2%;  Score 62;  DB 4;  Length 978;
  Best Local Similarity   47.1%;  Pred. No. 0.00029;
  Matches  225;  Conservative    0;  Mismatches  250;  Indels    3;  Gaps    1;
```

```
Qy          96 TAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTAC 155
               |||   || |   ||      ||| |||  |||||  |||||||||  | | ||||   ||
Db         117 TAATCAAATTGTTGAACAGCAAAATGATAAAGTTGTTATCAAGCACCTATTAGGTGAAAC 176

Qy         156 GAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTT 215
                  |||| ||||| | ||       |||||  |  |||| | |    |   |||  | ||
Db         177 TGCAGTATCTAAAAATCCGTCTAAGGTTGTACTGTTTGATTTTGGTCTATATGACTCTTT 236

Qy         216 AGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTAT 275
               |||||| |||| | |   | | | |    | |  | |        |  ||||     |
Db         237 AGTTGCATTAGGTTT---AGCGGATAAAGTGGTTGGTTTACCATTAGGAAATGCACCTGA 293

Qy         276 TAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAA 335
                 |      | |   ||         |     ||  ||| ||   |   ||    ||| |
Db         294 ATATATTAAAGGGCGTATTGCGAATGATGTGGCTAATGTTGGGGGAATGAAAGCCCCTGA 353

Qy         336 CTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAA 395
                || ||   | ||  | | |||||||| ||||||||||||  |     |  |  |  |
Db         354 TTTTGAGAGACTCGCTGAGCTTAAACCTGATTTAATTATTATTACCGGTCGTCAAGGCGC 413

Qy         396 AGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGG 455
                  |  |   |||  |  ||| |      |  | |  | |  |   | ||  |   |
Db         414 TTCTTATGAGAAACTCACAAACATTGCCCCCACGATTAATTTAGGTACAAATAGTGCGAA 473

Qy         456 AGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGA 515
                 ||  | ||  |      |    | || ||  |||     ||   || | |||  || |
Db         474 TTATCTTGATTCAGTGGAAGCGAATATTCAATTAATAGGTGAATTATTTGATAAGCAACA 533

Qy         516 AGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAA 573
               ||    | ||| ||  |||| |  |    |||     |||||| |   | | ||| ||
Db         534 AGCTACTCAAACACAATTAGCAACATTGAATACCGTGATTGAACAGGCTCAGAAAAAA 591

RESULT 32
US-08-956-171E-37
; Sequence 37, Application US/08956171E
; Patent No. 6593114
;   GENERAL INFORMATION:
;        APPLICANT: Charles Kunsch
;                   Gil H. Choi
;                   Patrick S. Dillon
;                   Craig A. Rosen
;                   Steven C. Barash
;                   Michael R. Fannon
;        TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;        NUMBER OF SEQUENCES: 5256
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/956,171E
;             FILING DATE: 20-Oct-1997
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/009,861
;             FILING DATE: January 5, 1996
;             APPLICATION NUMBER: 08/781,986
;             FILING DATE: January 3, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mark J. Hyman
;             REGISTRATION NUMBER: 46,789
;             REFERENCE/DOCKET NUMBER: PB248P1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (240) 314-1224
;             TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 37:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9834 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 37:
US-08-956-171E-37

  Query Match             6.1%;  Score 61.8;  DB 4;  Length 9834;
  Best Local Similarity   47.0%;  Pred. No. 0.00052;
  Matches  334;  Conservative    0;  Mismatches  362;  Indels   15;  Gaps    4;

Qy         291 TAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAG 350
               ||||  ||  ||||| ||    || || |     ||  ||   || || ||  ||||
Db          28 TAAAGATGATAAATATACAAACGTTGGTAATTTAAAAGAAGTGAATTTTGATAAAATTGC 87

Qy         351 TAAACTTAAACCAGATTTAA---TTATTGCTGATAATAATAGACACAAAGGTATTTATAA 407
               |      ||||| ||  |||   ||||  |||    ||      |  ||   | ||   |
Db          88 TGCGACGAAACCCGAAGTAATCTTTATCTCTGGACGTACAGCTAATCAAAAGAATTTAGA 147

Qy         408 AGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGA 467
                 | || || |||  ||| ||||      || |      | | | || |||| | ||  |
Db         148 TGAATTCAAAAAAGCTGCACCTAAAGCGAAAATTGTTTATGTTGGTGCAGATGAAAAGAA 207

Qy         468 AAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAA 527
                   |||| | |  | |||  ||   |  ||  | | || |||    | ||   |||| |
Db         208 CTTAATTGGTTCAATGAAACAAAACACTGAAAATATCGGAAAAATTTACGA---TAAAGA 264

Qy         528 ACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAA 587
               |     ||   || |    || |||  |  |||  | ||||  |    ||  |  |||||
Db         265 AGATAAAGCTAAAGAATTAAATAAAGATTTAGATAACAAAATTGCTTCAATGAAAGATAA 324

Qy         588 AAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAA 647
               ||   ||||  |        | ||      |  |  || |      | | |   |   |
Db         325 AACGAAAAACTTCAATAAAACTGTTATGTATTTACTAGTTAACGAAGGTGAATTATCAAC 384

Qy         648 CTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGT 707
                | |     |          |  ||  | ||| ||   |   |  |  |    |||  ||
Db         385 ATTTGGACCTAAAGGTCGTTTTGGTGGATTAGTTTACGATACATTAGGATTCAATGCAGT 444

Qy         708 TACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTT 767
               |  ||||     |||||||      | | ||||       | |  |  |    |  |||  | |
Db         445 TGATAAAAAAGTAAGTAATAGCAATCATGGACA------AAATGTTTCTAACGAATATGT 498

Qy         768 ATCTCAAGTGAATCCTGAGCGTATGTTCATAATG---ACAAACAAAGCAAGTTCTAACGA 824
                  | |||  ||||| ||   ||| ||    |||   | |    |||| |    |    |
Db         499 TAATAAAGAAAATCCAGATGTTATTTTAGCGATGGATAGAGGTCAAGCGATAAGTGGTAA 558

Qy         825 ACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAA 884
               | |  |    ||| ||  |  |||  |||||||||  || ||  | || ||  | ||| |
Db         559 ATCAACTGCGAAACAAGCATTAAATAATCCTGTATTAAAAAATGTTAAAGCAATTAAAGA 618

Qy         885 TCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTC 944
                 |    ||  ||| ||||| | | | ||||||       ||  ||| | || | |  |
Db         619 AGACAAAGTATATAATTTAGATCCTAAATTATGGTACTTTGCAGCTGGATCAACTACAAC 678

Qy         945 AGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAA 995
                  |   |    ||  ||     ||||   |  |   ||| |   | |||||
Db         679 TACAATTAAACAAATTGAGGAACTTGATAAAGTTGTAAAATAATTTTAAAA 729

RESULT 33
US-08-781-986A-37
; Sequence 37, Application US/08781986A
; Patent No. 6737248
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  5255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;       STREET:  9410 Key West Avenue
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  USA
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;       COMPUTER:  HP Vectra 486/33
;       OPERATING SYSTEM:  MSDOS version 6.2
;       SOFTWARE:  ASCII Text
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/781,986A
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Benson, Bob
;       REGISTRATION NUMBER:  30,446
;       REFERENCE/DOCKET NUMBER:  PB248PP
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 309-8504
;       TELEFAX:  (301) 309-8512
;   INFORMATION FOR SEQ ID NO:  37:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9834 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
US-08-781-986A-37

  Query Match             6.1%;  Score 61.8;  DB 4;  Length 9834;
  Best Local Similarity   47.0%;  Pred. No. 0.00052;
  Matches  334;  Conservative    0;  Mismatches  362;  Indels   15;  Gaps    4;

Qy        291 TAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAG 350
              ||||  ||  ||||| ||    || || |     ||  ||   || || ||  ||||
Db         28 TAAAGATGATAAATATACAAACGTTGGTAATTTAAAAGAAGTGAATTTTGATAAAATTGC 87

Qy        351 TAAACTTAAACCAGATTTAA---TTATTGCTGATAATAATAGACACAAAGGTATTTATAA 407
              |      ||||| ||  |||   ||||  |||    ||      |  ||   | ||   |
Db         88 TGCGACGAAACCCGAAGTAATCTTTATCTCTGGACGTACAGCTAATCAAAAGAATTTAGA 147

Qy        408 AGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGA 467
               || || || |||  ||| ||||      || |      | | | || ||||| | ||  |
Db        148 TGAATTCAAAAAAGCTGCACCTAAAGCGAAAATTGTTTATGTTGGTGCAGATGAAAAGAA 207

Qy        468 AAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAA 527
                  |||| | |  | ||| ||   |  ||  | | || |||    | ||   |||| |
Db        208 CTTAATTGGTTCAATGAAACAAAACACTGAAAATATCGGAAAAATTTACGA---TAAAGA 264

Qy        528 ACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAA 587
              |     ||   || |    || |||  |  |||  | |||||  |    ||  |  |||||
Db        265 AGATAAAGCTAAAGAATTAAATAAAGATTTAGATAACAAAATTGCTTCAATGAAAGATAA 324

Qy        588 AAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAA 647
              ||   ||||  |        | ||      |  |  || |      | | |   |  ||
Db        325 AACGAAAAACTTCAATAAAACTGTTATGTATTTACTAGTTAACGAAGGTGAATTATCAAC 384

Qy        648 CTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGT 707
               | |     |          |  ||  | ||| ||   |   |  |  |  |    |||  ||
Db        385 ATTTGGACCTAAAGGTCGTTTTGGTGGATTAGTTTACGATACATTAGGATTCAATGCAGT 444

Qy        708 TACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTT 767
              |  ||||    |||||||      | | ||||       | |  |    |  |||  | |
Db        445 TGATAAAAAAGTAAGTAATAGCAATCATGGACA------AAATGTTTCTAACGAATATGT 498

Qy        768 ATCTCAAGTGAATCCTGAGCGTATGTTCATAATG---ACAAACAAAGCAAGTTCTAACGA 824
                 | |||  ||||| ||   ||| ||    |||   | |    |||| |    |    |
Db        499 TAATAAAGAAAATCCAGATGTTATTTTAGCGATGGATAGAGGTCAAGCGATAAGTGGTAA 558

Qy        825 ACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAA 884
              | |  |    ||| ||  |  |||  ||||||||||  || ||  | || ||  | ||| |
Db        559 ATCAACTGCGAAACAAGCATTAAATAATCCTGTATTAAAAAATGTTAAAGCAATTAAAGA 618

Qy        885 TCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTC 944
                |    ||  ||| |||||| | | | ||||||       ||  ||| | || | |  |
Db        619 AGACAAAGTATATAATTTAGATCCTAAATTATGGTACTTTGCAGCTGGATCAACTACAAC 678

Qy        945 AGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAA 995
                 |   |    ||  ||     ||||   |  |   ||| |   | ||||
Db        679 TACAATTAAACAAATTGAGGAACTTGATAAAGTTGTAAAATAATTTTAAAA 729

RESULT 34
US-09-601-198-75/c
; Sequence 75, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
;  APPLICANT: Cassell, Gail H.
;  APPLICANT:  Chen, Ellson Y.
;  APPLICANT:  Glass, Jennifer S.
;  APPLICANT:  Glass, John I.
;  APPLICANT:  Heiner, Cheryl R.
;  APPLICANT:  Lefkowitz, Elliot
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
;  TITLE OF INVENTION:  UREALYTICUM
;  FILE REFERENCE: UAB-13452/22
;  CURRENT APPLICATION NUMBER: US/09/601,198
;  CURRENT FILING DATE:  2000-12-08
;  PRIOR APPLICATION NUMBER: 60/073,189
;  PRIOR FILING DATE: 1998-01-30
;  NUMBER OF SEQ ID NOS: 181
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
;   LENGTH: 5652
;   TYPE: DNA
;   ORGANISM: Ureaplasma urealyticum
US-09-601-198-75

  Query Match             5.9%;  Score 59.8;  DB 4;  Length 5652;
  Best Local Similarity   44.4%;  Pred. No. 0.0013;
  Matches  241;  Conservative    0;  Mismatches  302;  Indels    0;  Gaps    0;

Qy        225 AGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATT 284
              | |  ||||   | ||   | ||  | || | |  | ||| || |||  | || |
Db       3853 AAAAATTAATGGTATTCATACAGAAGTTGTTCAAGATAAATATGGTAAAAATACATATAA 3794

Qy        285 AAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGA 344
              |   | |||   ||  | | |   |  |   | |||   |      ||| | |||  |
Db       3793 AGTTGTTAATGGTGAATACGAAGGTATTTATGTAACTGATCGAATTCCTTATTTATCATT 3734

Qy        345 AATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTA 404
              | |   | ||  | | || |  || |  |   | | || |||       |||  | |
Db       3733 AGTTGCTCAATCTGATCCTGCATTTAAAACAACAGGTATTAACTATTTAAAATATGTAAG 3674

Qy        405 TAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAA 464
              || | ||  | ||  |  |   |  ||  |  ||   |  |       | | |||  |||
Db       3673 TACACACGAATATGGACATCACCAAACATTACAAGACATGAAAGATATTAGTGATAGTAA 3614

Qy        465 TGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAA 524
              |||     ||  || |     ||    ||| | | |   || ||  || |||  |  |
Db       3613 TGATTCAGTTATTGGTGGAGGAATTGATTCTAGATCGGGAGTTAGTGATGAATCATATGT 3554

Qy        525 AAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGA 584
              |||  |    | |  ||| ||| |   || || |   | ||    |   |    ||   |
Db       3553 AAATGGTAAGGCATTACAAGATTATTTAAATGCACGTTCTAGCGGAATTACCTTTAGAAA 3494

Qy        585 TAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAG 644
               |   ||     |       ||  ||  || ||    ||    |    ||  |       |
Db       3493 AACTGATGTTGATTATAATCCAACAAAAGATGGGTCATTCTTTAATTTTTCATTAAATAA 3434

Qy        645 CAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGA 704
              | |  || |   |      |    ||  |||  || | ||   |   |  | | || |||
Db       3433 CGATCCTAAAAATCCAGTTTGAGAAACACAAAAAGATATTTTTGGTTCCGTTAATGCTGA 3374
```

```
Qy           705 TGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAAC 764
                 || | ||||||  ||     |  ||   ||| | | ||  |||||||   ||  ||||
Db          3373 TGATCCTAAAGCCTTTTTCTATAATAAAAAACGTCGTTTTTTACAAAAATACGATGAATT 3314

Qy           765 TTT 767
                  ||
Db          3313 ATT 3311


RESULT 35
US-08-487-826B-13
; Sequence 13, Application US/08487826B
; Patent No. 5993827
;  GENERAL INFORMATION:
;    APPLICANT:  Sim, Kim L.
;    APPLICANT:  Chitnis, Chetan
;    APPLICANT:  Miller, Louis H.
;    APPLICANT:  Peterson, David S.
;    APPLICANT:  Su, Xin-zhaun
;    APPLICANT:  Wellems, Thomas E.
;    TITLE OF INVENTION:  BINDING DOMAINS FROM PLASMODIUM VIVAX
;    TITLE OF INVENTION:  AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe Martens Olson & Bear
;      STREET:  620 Newport Center Drive  16th Floor
;      CITY:  Newport Beach
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  92660
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,826B
;      FILING DATE:  10-SEP-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Israelsen, Ned
;      REGISTRATION NUMBER:  29,655
;      REFERENCE/DOCKET NUMBER:  NIH121.001CP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 235-8550
;      TELEFAX:  (619) 235-0176
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19124 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-487-826B-13

  Query Match             5.9%;  Score 59;  DB 2;  Length 19124;
  Best Local Similarity   50.1%;  Pred. No. 0.0024;
  Matches  173;  Conservative    0;  Mismatches  170;  Indels    2;  Gaps    1;

Qy           250 GATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACT 309
                 ||  | || |  |||||   ||| |  |||  ||||| | | ||  ||  | || | |
Db         15570 GAATAAAAAATGAAAAAGATTATCAAAAAAAAAATTAAAAAAAAATTTTATATAAAAAAAA 15629

Qy           310 TCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTA 369
                   ||     | |   ||  ||   ||   ||||||||  |  |||| | ||  | |   |
Db         15630 AATGATTATAAAAAAAATAAAAACAA--AAGAAGAAAAAAAAAAACATTAAAAAAAAAAA 15687

Qy           370 ATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCT 429
                 | |||   | |||| || | | | ||| | |   |  |      ||  |||||   |
Db         15688 AATATATATCATAAAAACAAAAAAAAAAAGAAAAAAATATATTAAAATAAAAATATATATC 15747
```

```
Qy           430 ACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACA 489
                 |  |   ||   |||| |    |        | ||  || ||||  ||    | |||| |
Db         15748 ATAAAATAAAAAAAAAATTAAAAAAAATGTTAAAAAAAAAAATATATACATAAAATAAAAAAA 15807

Qy           490 ATTTCAAAAGCTTTAGGTAAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAG 549
                 ||||    |  |  |   ||| || ||  |   |||      | |  | || |  ||||
Db         15808 ATTTATTTAAATAAAAAAAAATAATAAATAAAAAAATTTAATTAAATAAAAAAAAATAAT 15867

Qy           550 AAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAA 594
                 |||    || | |  |  ||| ||| ||  ||  | |||| | ||
Db         15868 AAATAAAAAAATTTAATTAAATAAAAAAAAAATTAAAAAAATTTAA 15912


RESULT 36
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
;  APPLICANT: SHIGENOBU, SHUJI
;  APPLICANT:  WATANABE, HIDEMI
;  APPLICANT:  HATTORI, MASAHIRA
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
;  FILE REFERENCE: 081356/0159
;  CURRENT APPLICATION NUMBER: US/09/790,988
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: JP2000-107160
;  PRIOR FILING DATE: 2000-04-07
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 640681
;   TYPE: DNA
;   ORGANISM: Buchnera sp.
US-09-790-988-1

  Query Match             5.9%;  Score 59;  DB 4;  Length 640681;
  Best Local Similarity   49.6%;  Pred. No. 0.005;
  Matches  178;  Conservative    0;  Mismatches  180;  Indels    1;  Gaps    1;

Qy           255 TAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGT 314
                 ||| |||| |||   | ||||||||  || | ||  |   | |   | ||     | | |
Db        635938 TAAAAAAAGAAACTCTTTTATTAAAAAATAATGATTTTTTAATTAGATATTTTTCTTTAT 635997

Qy           315 AGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTAT 374
                 | |||||  ||   ||  |||   |||| || |     ||   | |   ||   |  |
Db        635998 ATGAACAATTAGATAATTTAAAAGAGAATAATTTTAATAAAAAAGAATTGAACAATCTCG 636057

Qy           375 TGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGAT 434
                   |  ||||||  | | | ||   ||| |   || |   ||||| ||||     |   |
Db        636058 AACAAATAATACAAAAAAAAATTATATATTGCAAAAAAAAAATAGAATTAGAGTTTGAAA 636117

Qy           435 TGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTC 494
                   |   || |        |    |  |||||   |  |||  |  ||||  || || |
Db        636118 AAACAAGATATAAAAAAGTCATTAAAATAATTTCAGAATTACTTTTTTTTGAAAAAATAA 636177

Qy           495 AAAAGCTTTAGGTAAAGAAGAAGGTA-AAAAACGCTTAGAAGAACACGATAAGAAAA 553
                 || |  ||| |  |||  |  | ||  || |   | ||  |  | |  | ||||     |
Db        636178 AAGACGTTTTGAAAAAAGAATACAATATATATTTAAGCCAAATAAATTAGGATATTTTTA 636237

Qy           554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTA 612
                 |   |   | |  |||| |||  || ||  | ||  || |  ||||||||  |   |||
Db        636238 TATGATTTTTTTAAAAAAAAACACGATAAAAAATTATTATTAGGTATTGATTTAGGTA 636296


RESULT 37
US-09-601-198-66
; Sequence 66, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
;  APPLICANT: Cassell, Gail H.
;  APPLICANT:  Chen, Ellson Y.
; APPLICANT:  Glass, Jennifer S.
; APPLICANT:  Glass, John I.
; APPLICANT:  Heiner, Cheryl R.
; APPLICANT:  Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; TITLE OF INVENTION:  UREALYTICUM
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE:  2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66
;   LENGTH: 861
;   TYPE: DNA
;   ORGANISM: Ureaplasma urealyticum
US-09-601-198-66

  Query Match              5.8%;  Score 58.8;  DB 4;  Length 861;
  Best Local Similarity   44.8%;  Pred. No. 0.0014;
  Matches  225;  Conservative    0;  Mismatches  277;  Indels    0;  Gaps    0;

Qy        251 ATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTT 310
              | | |||  |||| |     |||||| |   |  | | | || |||    ||| | ||
Db        206 AGGTTAATGAAAATATAACAATTATTGATAAAAAAGGTGGTATAATCCATAAAAAGACAG 265

Qy        311 CTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAA 370
                 |   || |  ||| |||   ||   | || ||   | |  |    |         ||
Db        266 ACATTACAAAAATTAACCAAATAAATCAACAAAAAGAAATTTTAAAGGATTATTTAAAAA 325

Qy        371 TTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTA 430
               ||  |  || || ||||   |  ||   | | |||   ||  ||||  || |   || |
Db        326 ATAAAGAAGAAAAAAAATAAGGATCAAATAAATAATACTAACGAAAATCTAAATAAACCAA 385

Qy        431 CGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAA 490
                ||| |  | |||| | | ||||   |  | ||  || | | |    ||     || ||
Db        386 TAATTAATGTTAAAAATGTAGATGATAAAAACAAACAAGAAAATTCAACTAAATTAAAAA 445

Qy        491 TTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGA 550
               |    |   | |   |||  | ||  ||   || |||  |   ||| |  |  |||  |
Db        446 ATAATGATTTTATTTCTAATAATGATAAAAACAATAAAATCAACGAAAATAATAATATAA 505

Qy        551 AAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAG 610
                  ||||||| |        || |||        || | |   || |||  |  |   |
Db        506 GTTATGAAGAAAAACCTTTTAAACTAAAACGCTTCAATATTATTAATGTAATTGCGAAAA 565

Qy        611 TAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAA 670
                   | |  ||     ||  | | |     ||   | | || |  ||   |  |    |
Db        566 ATAAAGATTTATACCAATTATTAGATTTCAAAACTTATTTTTCTAATGAAAACATTAATA 625

Qy        671 GTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATC 730
               | |  |    ||||| |||   ||   |  | |   ||   | |   ||   || | |
Db        626 ATGATATTAAATTTAATGAAAAGTTGTTTATTAAAAATATCTATGAAATAGTCAAATCTG 685

Qy        731 TTAAAGGACCTTACTTACAAAT 752
               || | |   ||    | || |
Db        686 CTATAAGTAGTTTTCAAGAATT 707

RESULT 38
US-09-949-016-14624/c
; Sequence 14624, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14624
;   LENGTH: 147382
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(147382)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14624

  Query Match              5.6%;  Score 56.8;  DB 4;  Length 147382;
  Best Local Similarity   48.2%;  Pred. No. 0.011;
  Matches  193;  Conservative    0;  Mismatches  202;  Indels    5;  Gaps    1;

Qy        196 TATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGAT 255
              |||  |  | |  ||   | | |   | || || |  ||     |   |      || |
Db     137260 TATATAAATATATATAAATAAATAAATATAAATATATAAATATATATAAATATATATAAA 137201

Qy        256 AACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTA 315
               | | ||| |    |||    | |    |||   |  ||||   || |||  |  | ||||
Db     137200 TAAATAAATATAAATATATAAATATATATAAATAAATAAATATAAATATATAAATATGTA 137141

Qy        316 GGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATT 375
                || || ||   ||   ||  || |   |   |   |  | ||  || || ||| |||
Db     137140 TAAATACATATATAAATAAATATAAATATATAAATATATATAAATACATATATAAATATA 137081

Qy        376 GCTGATAATA-----ATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTA 430
                | |  |||       | | ||| |  ||| |||||| |  |||| | |  |     ||
Db     137080 TATAAATATATATATTAATATACATAAATATGTATAAATATATAAACATATATTAATATA 137021

Qy        431 CGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAA 490
                | |  |   |||  | |  ||  | ||  | ||  | | ||  |||  | || |  ||
Db     137020 TAAATATATATAAATATATTAATATATATATTTATATATAAATATATGAATATATATAAA 136961

Qy        491 TTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGA 550
              | |   ||  | ||  || | |  || ||    | |||    || |  || |     |
Db     136960 TATGTTAATATATACATATATATAAATAAATATATAAAAATATATATAAATATATATAAC 136901

Qy        551 AAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAA 590
               |||  |  |||||| | ||  ||||  |||  ||| | |
Db     136900 TAATATATAAATATATATAACTAATATATATAAATATATA 136861

RESULT 39
US-09-902-540-1318
; Sequence 1318, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1318
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(614)
;   OTHER INFORMATION: unsure at all n locations
```

```
US-09-902-540-1318

  Query Match             5.6%;  Score 56.4;  DB 4;  Length 614;
  Best Local Similarity   47.7%;  Pred. No. 0.0042;
  Matches  165;  Conservative    0;  Mismatches  181;  Indels    0;  Gaps    0;

Qy        251 ATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTT 310
              || | || |||||||| |  |  |  |||  |  |||||| ||||    |||  | |
Db        180 ATAAAAAAAAAAAAAAACAAAAAAAAAAAAAAAAAAGAGAAAAAACAACAAACAAGAAAA 239

Qy        311 CTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAA 370
                |  | |  |   ||  ||    ||    || ||   |  |||    ||  |     ||
Db        240 AAGCCGAATAAAAAAACAAAATACACAATAAACAACAAAAAAAAGAACAAAAAACAACAA 299

Qy        371 TTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTA 430
                |      | ||  | ||||| |||   |   | | | |   |||  |||        |
Db        300 AAAGAAAAAAAAAAGAAAGACAAAAAAAAAACAAAACAAAAACAAAATAAAAAAAAAAAA 359

Qy        431 CGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAA 490
              | |   |     ||       |   | |  | ||  |||| |   |       |||| ||
Db        360 CAAAAAACACACAACAAAAAAACATAAAACAAAAAAAAAAAAGAAAAAAAAAAAAAAAAA 419

Qy        491 TTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGA 550
                  ||||     |   ||| || || ||   ||||||     | || ||    | ||||
Db        420 AAAAAAAAAAAAAAAAAAAAAAAAAAAAACAAAAAAAAAAAAAAAAAAAAACAAAAAAGA 479

Qy        551 AAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAA 596
              |||   || ||   || ||| ||| || ||   | |||||  ||||
Db        480 AAAAAAAAAAAACAAACAAAAAAAAAAATAAAAAAAAAAAAAAAAA 525

RESULT 40
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
;  APPLICANT: Bult et al.
;  TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ
; Patent No. 6503729
;  TITLE OF INVENTION:  jannaschii
;  FILE REFERENCE: PB275
;  CURRENT APPLICATION NUMBER: US/08/916,421B
;  CURRENT FILING DATE:  1997-08-22
;  PRIOR APPLICATION NUMBER: US 60/024,428
;  PRIOR FILING DATE: 1996-08-22
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1664976
;   TYPE: DNA
;   ORGANISM: Methanococcus jannaschii
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (28222)..(28222)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (28257)..(28258)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (84773)..(84773)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (84808)..(84808)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (84812)..(84812)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98120)..(98120)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98159)..(98159)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98239)..(98239)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98266)..(98266)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98343)..(98343)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (103998)..(103998)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (148948)..(148948)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (163385)..(163385)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (191989)..(191989)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (191995)..(191995)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (231980)..(231980)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (234187)..(234187)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (234220)..(234220)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (234814)..(234814)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (309398)..(309398)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (309418)..(309418)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (312837)..(312837)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (312993)..(312993)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (319226)..(319226)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (559167)..(559167)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (559241)..(559241)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (600992)..(600992)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (622708)..(622708)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (657081)..(657081)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (657203)..(657203)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (674435)..(674435)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (682442)..(682442)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (713652)..(713652)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (741684)..(741684)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (779455)..(779455)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (779676)..(779676)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (855539)..(855539)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (871619)..(871619)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1084830)..(1084830)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1096846)..(1096846)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1119881)..(1119881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1130881)..(1130881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1310988)..(1310988)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1313224)..(1313224)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1349473)..(1349473)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1349491)..(1349491)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1470091)..(1470091)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1569020)..(1569020)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1602912)..(1602912)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1603734)..(1603734)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1637998)..(1637998)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1664854)..(1664855)
;   OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

  Query Match             5.6%;  Score 56.4;  DB 4;  Length 1664976;
  Best Local Similarity   47.7%;  Pred. No. 0.022;
  Matches  204;  Conservative    0;  Mismatches  216;  Indels    8;  Gaps    1;

Qy        314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
              ||  || |   ||| ||        |  | | |  |  |||   |||    | |||   |
Db     661285 TATTAAAAGAAAAGAAAAAATTGGAAAGAAAGAAGAAAAAAGAGAAAGAGAAGTTAGAGA 661226

Qy        374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                |   | || ||   | | |||| ||   | ||| | | ||||   |||   |   | |
Db     661225 AAGAAAAGAAGAAAGAAGAAAAAGCTAAGGAAAAACAGTCAAATCTGATTATACAACCTA 661166

Qy        434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                ||| | |||      |    || | |    || || |  || |   |||||  ||
Db     661165 AAGAAATTAAAGAAGAGGTAAAAGCTGAAGTAGAGAAAAAAGAAGAAGTTAAAGAAAAGA 661106

Qy        494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
               |  ||    |   ||||  |||||||  ||| ||   | | ||  | | || ||  |||
Db     661105 TAGTAGAAAAACCAAAAGCTGAAGAAGTAAAAGAAAAATCAAAAACAGAAGAAAAAGAAA 661046

Qy        554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGC--------C 605
               | || || |||||||  |   ||  |   | ||| |  | || | |  ||
Db     661045 CTAAAAAAGATAAAAAGAAGGGAAAGAAAAAGAAAGAAGATAAAGGAAAGCAACTAACTT 660986

Qy        606 TGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATT 665
              |  |     |  | ||  |  |||  ||    | |  | | | | |||  |  |    ||
Db     660985 TAGATGCTTTCTTCAAATAAATTTATTTAAAAAACTTATCTAATGTTAATTGAGCTCCTT 660926

Qy        666 CCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAA 725
              | |   ||||||   ||||  | |  | |   |     ||  | || ||   |  || ||
Db     660925 CTTTCTTCAACTCATTTTTTGAAACTCCTACGGCTTCCATAATTCTCAAAACTGGAGGAA 660866

Qy        726 GTATCTTA 733
              | |||| |
Db     660865 GAATCTGA 660858

Search completed: August 26, 2005, 07:02:59
Job time : 220 secs
```

This Page Blank (uspto)

2

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2005  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         August 26, 2005, 00:31:21 ; Search time 4629 Seconds
                                           (without alignments)
                                           10551.492 Million cell updates/sec

Title:          US-10-724-972A-2580
Perfect score:  1008
Sequence:       1 ggagtggaatcagtgagagg..........gtaaaaaagataataagtaa 1008

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters:      9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       GenEmbl:*
                1:  gb_ba:*
                2:  gb_htg:*
                3:  gb_in:*
                4:  gb_om:*
                5:  gb_ov:*
                6:  gb_pat:*
                7:  gb_ph:*
                8:  gb_pl:*
                9:  gb_pr:*
                10:  gb_ro:*
                11:  gb_sts:*
                12:  gb_sy:*
                13:  gb_un:*
                14:  gb_vi:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES

                       %
Result               Query
   No.     Score   Match Length DB  ID                   Description
-------------------------------------------------------------------------------
c   1       1008   100.0 300698  1  AE016750             AE016750 Staphyloc
c   2        569    56.4 110000  1  BX571857_22          Continuation (23 o
c   3        569    56.4 304050  1  AP004829             AP004829 Staphyloc
c   4        569    56.4 307750  1  AP003136             AP003136 Staphyloc
c   5        569    56.4 347650  1  AP003364             AP003364 Staphyloc
c   6      567.4    56.3 110000  1  BX571856_23          Continuation (24 o
    7      557.4    55.3   1044  6  AX622368             AX622368 Sequence
c   8      336.2    33.4   2247  6  A64853               A64853 Sequence 10
    9      269.2    26.7 298050  1  AP003132             AP003132 Staphyloc
   10      269.2    26.7 347785  1  AP003361             AP003361 Staphyloc
   11      263.8    26.2    975  6  AX621226             AX621226 Sequence
   12      263.2    26.1    957  6  AX583924             AX583924 Sequence
   13      262.8    26.1 110000  1  BX571857_10          Continuation (11 o
   14      262.8    26.1 290150  1  AP004825             AP004825 Staphyloc
   15      254.2    25.2 110000  1  BX571856_10          Continuation (11 o
   16      222.6    22.1    668  6  AR354237             AR354237 Sequence
   17      222.6    22.1    668  6  AR535793             AR535793 Sequence
c  18      190.8    18.9   5718  6  AX951891             AX951891 Sequence
   19      190.8    18.9  37900  1  D86417               D86417 Bacillus su
```

```
c  20      190.8    18.9 197409  1  BSUB0005             Z99108 Bacillus su
   21      168.4    16.7 110000  1  AE017333_11          Continuation (12 o
   22      168.4    16.7 110000  1  CP000002_11          Continuation (12 o
   23      166.6    16.5  12556  1  BSY14084             Y14084 Bacillus su
   24      166.6    16.5 198743  1  BSUB0006             Z99109 Bacillus su
   25      162.2    16.1    242  6  AR356438             AR356438 Sequence
   26      162.2    16.1    242  6  AR537994             AR537994 Sequence
c  27      156.6    15.5  13792  1  AE006047             AE006047 Pasteurel
   28      149.4    14.8    801  6  AX433651             AX433651 Sequence
   29      146.4    14.5 301332  1  AE017012             AE017012 Bacillus
   30      143.2    14.2 110000  1  AE017225_43          Continuation (44 o
   31      143.2    14.2 110000  1  AE017334_43          Continuation (44 o
   32      143.2    14.2 290029  1  AE017038             AE017038 Bacillus
   33        140    13.9 110000  1  AE017355_43          Continuation (44 o
c  34      125.6    12.5 110000  1  BX571857_00          BX571857 Staphyloc
c  35      125.6    12.5 110000  1  BX571857_01          Continuation (2 of
c  36      125.6    12.5 290250  1  AP004822             AP004822 Staphyloc
   37      125.4    12.4   2115  6  AR354486             AR354486 Sequence
   38      125.4    12.4   2115  6  AR536042             AR536042 Sequence
c  39      124.6    12.4 110000  1  AE017355_06          Continuation (7 of
c  40      124.6    12.4 110000  1  CP000001_06          Continuation (7 of
   41      124.2    12.3 304262  1  AE017005             AE017005 Bacillus
   42        124    12.3    990  6  AX583941             AX583941 Sequence
   43        124    12.3    999  6  BD237137             BD237137 Staphyloc
   44        124    12.3    999  6  AR282355             AR282355 Sequence
   45        124    12.3   3180  1  AF079518             AF079518 Staphyloc

                              ALIGNMENTS

RESULT 1
AE016750/c
LOCUS       AE016750              300698 bp    DNA     linear   BCT 01-JAN-2003
DEFINITION  Staphylococcus epidermidis ATCC 12228, section 7 of 9 of the
            complete genome.
ACCESSION   AE016750 AE015929
VERSION     AE016750.1  GI:27316220
KEYWORDS    .
SOURCE      Staphylococcus epidermidis ATCC 12228
  ORGANISM  Staphylococcus epidermidis ATCC 12228
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1  (bases 1 to 300698)
  AUTHORS   Zhang,Y., Ren,S., Li,H., Fu,G., Lu,L., Lu,G., Jia,J., Tu,Y.,
            Qin,Z., Chen,Z. and Wen,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-NOV-2002) Chinese National Human Genome Center at
            Shanghai, 250 Bi Bo Road, Shanghai 201203, China
FEATURES             Location/Qualifiers
     source          1..300698
                     /organism="Staphylococcus epidermidis ATCC 12228"
                     /mol_type="genomic DNA"
                     /strain="ATCC 12228"
                     /db_xref="ATCC:12228"
                     /db_xref="taxon:176280"
     tRNA            complement(114..186)
                     /product="tRNA-Lys"
     tRNA            complement(190..264)
                     /product="tRNA-Gln"
     tRNA            complement(269..352)
                     /product="tRNA-Tyr"
     tRNA            complement(369..444)
                     /product="tRNA-Val"
     tRNA            complement(451..522)
                     /product="tRNA-Glu"
     tRNA            complement(524..598)
                     /product="tRNA-Asn"
     rRNA            complement(612..726)
                     /product="5S ribosomal RNA"
     rRNA            complement(886..3713)
                     /product="23S ribosomal RNA"
     gene            complement(2718..2831)
                     /locus_tag="SE1755"
```

```
CDS             complement(2718. .2831)
                /locus_tag="SE1755"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="AAO05396.1"
                /db_xref="GI:27316221"
                /translation="MMIIGGRALFGRGAPLGLPNSDKLRMPINLTWESEHG"
rRNA            complement(3977. .5530)
                /product="16S ribosomal RNA"
gene            complement(5985. .7052)
                /locus_tag="SE1756"
CDS             complement(5985. .7052)
                /locus_tag="SE1756"
                /codon_start=1
                /transl_table=11
                /product="ATP-binding Mrp-like protein"
                /protein_id="AAO05397.1"
                /db_xref="GI:27316222"
                /translation="MLTVNQVKEIVGALKDPIIDVPLRESEGIVDVSIKDNINHVSVK
                VAMAQLGGQPQLDLQMAIVKALKENGANTVGIRFEELPSEVVERYIGKGSEKPKTIEE
                LLSQNNPVEFISIASGKGGVGKSTVAVNLAVALAREGKKVGLVDADIYGFSVPDMMGI
                DERPGIDGKEIIPVERHGVKVISMAFFVEENAPVIWRGPMLGKMLTNFFTEVQWGELD
                YLLLDLPPGTGDVALDVHSMLPSSKEIIVTTPHPTAAFVAARAGAMAKHTEHTILGVI
                ENMSYFESKETGKKEYVFGKGGGKKLSDELETQLFAELPLEQPTWNPNDFSPSIYQSD
                DRLGELYTLIARKVIASTQKQ"
gene            complement(7203. .8642)
                /locus_tag="SE1757"
CDS             complement(7203. .8642)
                /locus_tag="SE1757"
                /codon_start=1
                /transl_table=11
                /product="multidrug resistance protein"
                /protein_id="AAO05398.1"
                /db_xref="GI:27316223"
                /translation="MASYELSKTKRNLIVSVMLISAFVAILNQTLLNTALPHIMRELN
                INENTSQWLITGFMLVNGVMIPLTAYLMDKVKTRPLYLGAMGSFLIGSIVAAIAPNFG
                VLMIARVIQAVGAGILMPLMQFTLFTLFSKEHRGFAMGLAGLVIQFAPAIGPTFTGLI
                IDNASWRVPFIVIVGIALVTFIFGFVSISSYNTTKETKLDKRSVIYSTLGFGLMLYAF
                SSAGNLGFSNPIVLCSVLISLIIIGIFVKRQITISNPLLNLKIFNNKIFCFSTITSMI
                IMLSMVGPALLIPLYVQNALGLSALLSGLVIMPGAIINGIMSVFTGKFYDKYGPRPLI
                FTGFILLISCTFLLCFLKVDTSYTYLIVIYAIRMFAVSLLMMPINTTGINALKTEDIS
                HGTAIMNFGRVMAGSLGTALMVTFMSIGAQWVVSSSEHASKEMIQRQSVAVGVDVSFA
                LVTVFVIIAFVLALFIKEPNHLTHNSRKV"
gene            complement(8912. .9382)
                /locus_tag="SE1758"
CDS             complement(8912. .9382)
                /locus_tag="SE1758"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="AAO05399.1"
                /db_xref="GI:27316224"
                /translation="MMKFLKNKSYHLLVTLIVLTIFVISGAIFLTFLGFGLYGLSRIL
                IYLHLGDFSYNKGFYDNLIYYGSYIVLGYFTLFSIEHLMDYFKKNLPKNPYFQGINFH
                LISYIVTTIMFYFIVHIHYVHVNIHFWVIMIIIGFLFVCKEVFYPESKNLNNKK"
gene            complement(9570. .10907)
                /locus_tag="SE1759"
CDS             complement(9570. .10907)
                /locus_tag="SE1759"
                /codon_start=1
                /transl_table=11
                /product="multidrug transporter"
                /protein_id="AAO05400.1"
                /db_xref="GI:27316225"
                /translation="MNLKSIITVMALILIMFMAAIETSIISLALPTIKNSLNAGNLVS
                LVFTVYFIALVIANPIVGELMSRFKIIYIAVVGVLLFALGSLMSGLSQTFTFLIISRT
                VQGFGAGVMMSLSQIVPKLAFEIPLRYKIMGIVGSVWGISSIIGPLLGGAILEFASWH
                WLFYINIPIAIVAIILVLMTFHFPDETQVQQSRFDIKGLIIFYIFIALLMFGLLNQHH
                IIFNIFSIILALAVLWLLFKIENSIEQPFLPTKEFNISIVLVFITDLLIAITLMGYNL
                YIPVYLQEKLSLSPLQSGFVIFPLSVAWITLNFNLAKIEAHFTRKTLYICSFFVLLVS
                SLMIMFGLKLPLLIAFAVVFAGLSFGYIYTKDSVIVQEETSPKNMKKMMSFYALTKNL
                GSSVGSTIMGYMYALNVGLFGSNLHNVLGLVLIIAVCLIVMWMTLYKSNTIQS"
```

```
gene            complement(11146. .11829)
                /locus_tag="SE1760"
CDS             complement(11146. .11829)
                /locus_tag="SE1760"
                /codon_start=1
                /transl_table=11
                /product="hemolysin III"
                /protein_id="AAO05401.1"
                /db_xref="GI:27316226"
                /translation="MSQSSKRKNDSVVETFKDIIPLTFGEEIGNAASHGAAALLTLFI
                LPYAAVHSFNNGGTLESISVSVYVISIFMMFISSTIYHSMQNNTSHKYILRIIDHSMI
                YVAISGTYTPVLLSVVGGWLGWLVTILLWGTTLWGILYKSIATKVNHRLSLIVYLVMG
                WVGIIFLPIIIMRTSWWFIFFIFLGGLSYTIGAWFYAQKNKPYFHMIWHIFIVLASFL
                HMIGIFYFM"
gene            complement(11848. .13035)
                /locus_tag="SE1761"
CDS             complement(11848. .13035)
                /locus_tag="SE1761"
                /codon_start=1
                /transl_table=11
                /product="UDP-N-acetylglucosamine pyrophosphorylase"
                /protein_id="AAO05402.1"
                /db_xref="GI:27316227"
                /translation="MLDKNQLEKYNQEHLYEYEKLMSSNEKNALDEKVDQLNLAEIQD
                LYQDLYVNRKTIDDVSSVSEVKYEVKSRLNEEERHTYEQKGYEAIRNGEFAVLLMAGG
                QGTRLGYKGPKGSFEIEGTSLFELQARQLIRLKEETGHTINWYIMTSDINHKDTIEYF
                KQHKYFNYDANHIHFFKQDNIVALSEEGKLVLNRDGHIMETPNGNGGVFKSLKKAGYL
                DKMQQDHVKYIFLNNIDNVLVKVLDPLFAGFTVTQSKDITSKTIQPKDSESVGRLVNV
                DCKDTVLEYSELDTDIVNQFNNANIGIHAFKLGFITSAVDRELPYHLAIKQLKQLDEN
                FGVVERPTLKFELFYFDIFRYGTSFVTLQVPREEEFSPLKNKEGKDSVHTATEDLKRM
                DLI"
gene            complement(13090. .13599)
                /locus_tag="SE1762"
CDS             complement(13090. .13599)
                /locus_tag="SE1762"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="AAO05403.1"
                /db_xref="GI:27316228"
                /translation="MTGKTHASCGFLVGAITTQYFHTDIFTSISVIVLSVISSILPDI
                CHTQSKIGRRFRLTSFFVRILFGHRTFTHSLLFIIGISFLLYFIQTPMYYMVAIVIGM
                FSHVILDILTPRGVKLLYPLPFNIVSPIHFKTGGLVDVSLATALSVGAIYTLFQPYLN
                TMMHYWLIK"
gene            13776. .14036
                /locus_tag="SE1763"
CDS             13776. .14036
                /locus_tag="SE1763"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="AAO05404.1"
                /db_xref="GI:27316229"
                /translation="MAMTVKKNDNEVRIQWRVADIKIPNNEIKNVTQDQDIHAVPEEN
                GKEISRIGSTFGKTNRVLIDTDQHLYIIYTQNDQKVYNELTK"
gene            14189. .14503
                /locus_tag="SE1764"
CDS             14189. .14503
                /locus_tag="SE1764"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="AAO05405.1"
                /db_xref="GI:27316230"
                /translation="MFVVTNRITVKKGYAKQMAPNFTKGGPIESLKGFEGIEVWQIDK
                DDYSEDMYVNSWWETEEDFKNWVNSDVFKQAHKNTGKSEDSPVIKSEIVKSNVLSSLN
                RR"
gene            complement(14682. .16055)
                /locus_tag="SE1765"
CDS             complement(14682. .16055)
                /locus_tag="SE1765"
                /codon_start=1
                /transl_table=11
```

```
                     /product="conserved hypothetical protein"
                     /protein_id="AAO05406.1"
                     /db_xref="GI:27316231"
                     /translation="MNGFKKVDIVRGRIKFITMSLLGIILFLVPIPVVQDGKQQTTLP
                     IAFLAGLLKDWLGGIMPILIVTIITVSGILTILCSTIYKNKLNPQGLMSSAFNVKIGW
                     LVLRVLAVFFSWLTFLNIGPEMIKSEDTGGLVFSSLLPTLVAVFLFAAIFLPLLMEYG
                     LLELLGPIFRPIMRPLFTLPGRSTVDNLASFIGDGTVGVLITSRQYGEGYYSRREATV
                     ISTTFSVVSITFAIVIAETIRMQDQFFYFYLTVVISCLIAAMIMPRIWPLKNIPDEYA
                     KEVSEEARNEQLPEGKTALKYGFDLATEVGIKSPGFKEFLISGFKTVVDMWFVILPVV
                     MSIGTIATIIANYTPVFEIIGKPFVPVLELLQIPEAHEASQTILIGFADMFLPSILIE
                     GVQNDVTRFVIGALSISQLVYLSEVGGVILGSKIPVSISKLFMIFLIRTIITLPIIAL
                     LAHLFIG"
     gene            complement(16352..17317)
                     /locus_tag="SE1766"

  Query Match             100.0%;  Score 1008;  DB 1;  Length 300698;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-144;
  Matches 1008;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGAGTGGAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      19364 GGAGTGGAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTA 19305

Qy         61 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      19304 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAA 19245

Qy        121 GATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      19244 GATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGT 19185

Qy        181 GTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      19184 GTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTT 19125

Qy        241 GGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      19124 GGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGA 19065

Qy        301 AAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      19064 AAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAA 19005

Qy        361 CCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      19004 CCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAA 18945

Qy        421 ATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18944 ATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCT 18885

Qy        481 TTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18884 TTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAA 18825

Qy        541 CACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18824 CACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTA 18765

Qy        601 TTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18764 TTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGT 18705

Qy        661 CAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18704 CAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTA 18645

Qy        721 AGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18644 AGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAAT 18585

Qy        781 CCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18584 CCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAA 18525

Qy        841 CTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18524 CTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATT 18465

Qy        901 TTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18464 TTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAA 18405

Qy        961 GAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db      18404 GAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 18357


RESULT 2
BX571857_22/c
WPCOMMENT
  Sequence split into 28 fragments  LOCUS BX571857 Accession BX571857
     Fragment Name       Begin      End
     BX571857_00             1   110000
     BX571857_01        100001   210000
     BX571857_02        200001   310000
     BX571857_03        300001   410000
     BX571857_04        400001   510000
     BX571857_05        500001   610000
     BX571857_06        600001   710000
     BX571857_07        700001   810000
     BX571857_08        800001   910000
     BX571857_09        900001  1010000
     BX571857_10       1000001  1110000
     BX571857_11       1100001  1210000
     BX571857_12       1200001  1310000
     BX571857_13       1300001  1410000
     BX571857_14       1400001  1510000
     BX571857_15       1500001  1610000
     BX571857_16       1600001  1710000
     BX571857_17       1700001  1810000
     BX571857_18       1800001  1910000
     BX571857_19       1900001  2010000
     BX571857_20       2000001  2110000
     BX571857_21       2100001  2210000
     BX571857_22       2200001  2310000
     BX571857_23       2300001  2410000
     BX571857_24       2400001  2510000
     BX571857_25       2500001  2610000
     BX571857_26       2600001  2710000
     BX571857_27       2700001  2799802
Continuation (23 of 28) of BX571857 from base 2200001 (BX571857 Staphylococcus aureus st

  Query Match             56.4%;  Score 569;  DB 1;  Length 110000;
  Best Local Similarity   73.3%;  Pred. No. 1.1e-77;
  Matches  728;  Conservative    0;  Mismatches  265;  Indels    0;  Gaps    0;

Qy          7 GAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCA 66
              | | |  |||||||| |||||| ||| ||| || | || ||| |  |||  ||| ||  |
Db      48085 GGAACGATGAGAGGTCTAAAAACTTTTAGTATATTGGGATTAATAGTTGCCTTACTTTTA 48026

Qy         67 ACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGA 126
                |||||| ||||| ||||    |      |    ||   ||   ||| | ||||||
Db      48025 GTTGCAGCTTGTGGTAATACGGATAATTCAAGTAAAAAAGAATCATCAACTAAAGATACT 47966

Qy        127 GTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTT 186
               ||    | ||  | ||| | |||||    |||||||||||| |  | |||||| | |||
Db      47965 ATTTCGGTAAAAGATGAAAATGGTACAGTAAAAGTACCTAAAGATGCAAAACGTATCGTT 47906

Qy        187 GTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATA 246
              ||  | ||||| |||||||  ||||| ||||  || ||||| ||||| || ||||| ||
Db      47905 GTATTAGAGTACTCATTTGCAGATGCATTAGCAGCATTAGACGTTAATCCAGTTGGTATT 47846

Qy        247 GCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATAC 306
              || ||||||   || ||||| |||| |||||||||| | ||||| ||||||||  | ||
Db      47845 GCTGATGATGGTAAGAAAAAACGTATCATTAAACCAGTTAGAGAAAAAATTGGGGATTAT 47786
```

```
Qy          307 ACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGAT 366
                ||||||||||| |||||||| || || |||||||||||||| |||||| | ||||| |||
Db        47785 ACTTCTGTAGGTACACGTAAACAGCCAAACTTAGAAGAAATTAGTAAATTAAAACCGGAT 47726

Qy          367 TTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCT 426
                |||||||| |||||||  | |||||| ||||||||| ||||||| ||||| ||||||||
Db        47725 TTAATTATCGCTGATAGCAGTAGACATAAAGGTATTAATAAAGAATTAAACAAAATTGCA 47666

Qy          427 CCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAA 486
                || ||  |   | | || ||||| |||||||| || || ||  |||| ||| || | || |||
Db        47665 CCAACATTATCATTAAAGAGTTTTGATGGAGACTACAAACAAAACATTAATTCGTTCAAA 47606

Qy          487 ACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGAT 546
                |||||| | |||||||||  ||||||| |||||||  |||||||  | |  || || |||
Db        47605 ACAATTGCTAAAGCTTTAAATAAAGAAAAAGAAGGCGAAAAACGTCTTGCTGAGCATGAT 47546

Qy          547 AAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCT 606
                ||   |||  |  | |||||| | ||||| |   | |||| ||||||||| ||  | ||
Db        47545 AAATTAATCAATAAGTATAAAGATGAAATTAAATTTGATAGAAATCAAAAAGTGCTTCCA 47486

Qy          607 GCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTC 666
                ||||||| ||||||| | |||||  | || ||||||| | | || |||||||| |||||
Db        47485 GCAGTAGTTGCTAAAGCTGGTTTATTAGCACATCCAAACTATTCATATGTTGGACAATTT 47426

Qy          667 CTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAG 726
                 |||   ||||||| |||||| | ||||||||||| ||||| || ||||||||||||||
Db        47425 TTAAACGAACTAGGATTTAAAAATGCATTAAGTGACGATGTAACAAAAGGTTTAAGTAAA 47366

Qy          727 TATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAG 786
                ||| | |||||||||||||||||| |  ||||||||  |||| || |  | ||||| |||
Db        47365 TATTTGAAAGGACCTTACTTACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAG 47306

Qy          787 CGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAA 846
                |||||| |||| |||||  |  | || |      |     ||  | | ||  || || ||
Db        47305 CGTATGATCATTATGACTGATCACGCTAAAAAAGATTCTGCTGAATTCAAGAAATTACAA 47246

Qy          847 AAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGAC 906
                 ||||| |   |||||| ||||| || || || |||||| | || || |||||| | |||
Db        47245 GAAGATGCAACATGGAAAAAAATTGAATGCAGTTAAAAATAATCGCGTGGATATTGTTGAC 47186

Qy          907 CGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTT 966
                |||||  | ||||||||||||| ||||| ||||||||||| ||||||||||| |||||||||
Db        47185 CGTGATGTTTGGGCAAGATCTCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTT 47126

Qy          967 GTTGAATTATCTAAGAAAGATAGTAAAAAAGAT 999
                ||||||||||| || |||||    ||  |||| |
Db        47125 GTTGAATTATCAAAAAAAGAACAAAAGTAAGGT 47093


RESULT 3
AP004829/c
LOCUS       AP004829               304050 bp    DNA     linear   BCT 20-DEC-2002
DEFINITION  Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,
            strain:MW2, section 8/10.
ACCESSION   AP004829 BA000033
VERSION     AP004829.1  GI:21205117
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus MW2
  ORGANISM  Staphylococcus aureus subsp. aureus MW2
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Baba,T., Takeuchi,F., Kuroda,M., Yuzawa,H., Aoki,K., Oguchi,A.,
            Nagai,Y., Iwama,N., Asano,K., Naimi,T., Kuroda,H., Cui,L.,
            Yamamoto,K. and Hiramatsu,K.
  TITLE     Genome and virulence determinants of high virulence
            community-acquired MRSA
  JOURNAL   Lancet 359 (9320), 1819-1827 (2002)
  MEDLINE   22040717
   PUBMED   12044378
REFERENCE   2  (bases 1 to 304050)
  AUTHORS   Director-General, Biotechnology Center, Aoki,K., Oguchi,A.,
            Nagai,Y., Asano,K., Iwama,N., Baba,T., Kuroda,M., Hiramatsu,K. and
```

```
            Kikuchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAR-2002) Director-General, Biotechnology Center,
            National Institute of Technology and Evaluation, Biotechnology
            Center; 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
            (E-mail:bio@nite.go.jp, URL:http://www.bio.nite.go.jp/,
            Tel:81-3-3481-1933, Fax:81-3-3481-8424)
FEATURES             Location/Qualifiers
     source          1. .304050
                     /organism="Staphylococcus aureus subsp. aureus MW2"
                     /mol_type="genomic DNA"
                     /strain="MW2"
                     /db_xref="taxon:196620"
     gene            complement(629. .1072)
                     /gene="MW1947"
     CDS             complement(629. .1072)
                     /gene="MW1947"
                     /note="ORFID:MW1947
                     hypothetical protein, similar to acetyltransferase
                     [Genomic island nu Sa 4mw]"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95812.1"
                     /db_xref="GI:21205118"
                     /translation="MSIITRLFNNSDFEKLNQLCKLYDDLGYPTNENDLKKRLKKITN
                     HDDYFLLLLIKENKIIGLSGMCKMMFYEKNAEYMRILAFVIHSEFRKKGYGKRLLADS
                     EEFSKRLNCKAITLNSGNRNERLSAHKLYSDNGYVSNTSGFTKQL"
     gene            complement(1178. .1696)
                     /gene="MW1948"
     CDS             complement(1178. .1696)
                     /gene="MW1948"
                     /note="ORFID:MW1948
                     [Genomic island nu Sa 4mw]"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical perotein"
                     /protein_id="BAB95813.1"
                     /db_xref="GI:21205119"
                     /translation="MDNINKNEKFRVLSTIAITISFILNFGLFISLAIWSINSMGNKQ
                     CISLNLLLSDDTFAVISVMTTLGTAILGGIAPSIYGIFTSPIDKKSSFRAISDWEQEQ
                     RNNVKDLMMLLPFTLLSILIVIFIFQFAYSKLFNTIAIIIFIIILVVMFFLILLICYK
                     ITVMICRYFSKE"
     gene            complement(1932. .2522)
                     /gene="MW1949"
     CDS             complement(1932. .2522)
                     /gene="MW1949"
                     /note="ORFID:MW1949
                     hypothetical protein, similar to bacteriophage terminase
                     small subunit [Genomic island nu Sa 4mw]"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95814.1"
                     /db_xref="GI:21205120"
                     /translation="MSELTAKQARFVNEYIRTLNVTQSAIKAGYNANSAHVTACRLLK
                     KPHIKQYIQEQKDKIIDENVLTAKELLHVLTNATVGDETETKEVVVKRGEYKENPQSG
                     KVQLVYNEHVELIEVPIKPSDRLKAREMLGKYHKLFTDKHDINGNVPIFINIGEWDGG
                     DEGLDKAVKDVSNDNPNHTVIVDDIPLEDYESIRFL"
     gene            complement(2519. .2734)
                     /gene="MW1950"
     CDS             complement(2519. .2734)
                     /gene="MW1950"
                     /note="ORFID:MW1950
                     [Genomic island nu Sa 4mw]"
                     /codon_start=1
                     /transl_table=11
                     /product="truncated hypothetical protein"
                     /protein_id="BAB95815.1"
                     /db_xref="GI:21205121"
                     /translation="MMTDDNVNDHIIKNHIEMIVDRLATDKEFYIFDSLIQGLSYQDI
                     SSALDCSEQSVILWYETILDKIVGVIK"
     gene            2864. .3338
                     /gene="truncated int"
     CDS             2864. .3211
```

```
                     /gene="truncated int"
                     /note="ORFID:MW1951
                     truncated hypothetical protein, similar to integrase
                     [Genomic island nu Sa 4mw]"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95816.1"
                     /db_xref="GI:21205122"
                     /translation="MKWSDIDFENNTISITKTYYNPNNNKKKFQILPPKTESSIGKIS
                     VDPNVIKVLRDYKINVQNNWKNELYNDNHCNEKTINVDTNHYVTYQHTQKYLYALIQI
                     HTLRIAHRIRCTH"
     CDS             3072. .3338
                     /gene="truncated int"
                     /note="ORFID:MW1952
                     truncated hypothetical protein, similar to integrase
                     [Genomic island nu Sa 4mw]"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95817.1"
                     /db_xref="GI:21205123"
                     /translation="MITIVMKKLSMWIQTIMSRTSILKNISTHSFRYTHCALLIESGV
                     HIKEIQERLRHKDINTTMNIYAKITNSYKKDASHKFSHQMEDVS"
     gene            complement(3433. .5049)
                     /gene="groEL"
     CDS             complement(3433. .5049)
                     /gene="groEL"
                     /note="ORFID:MW1953"
                     /codon_start=1
                     /transl_table=11
                     /product="GroEL protein"
                     /protein_id="BAB95818.1"
                     /db_xref="GI:21205124"
                     /translation="MVKQLKFSEDARQAMLRGVDQLANAVKVTIGPKGRNVVLDKEFT
                     APLITNDGVTIAKEIELEDPYENMGAKLVQEVANKTNEIAGDGTTTATVLAQAMIQEG
                     LKNVTSGANPVGLRQGIDKAVKVAVEALHENSQKVENKNEIAQVGAISAADEEIGRYI
                     SEAMEKVGNDGVITIEESNGLNTELEVVEGMQFDRGYQSPYMVTDSDKMVAELERPYI
                     LVTDKKISSFQDILPLLEQVVQSNRPILIVADEVEGDALTNIVLNRMRGTFTAVAVKA
                     PGFGDRRKAMLEDLAILTGAQVITDDLGLDLKDASIDMLGTASKVEVTKDNTTVVDGD
                     GDENSIDARVSQLKSQIEETESDFDREKLQERLAKLAGGVAVIKVGAASETELKERKL
                     RIEDALNSTRAAVEEGIVAGGGTALVNVYQKVSEIEAEGDIETGVNIVLKALTAPVRQ
                     IAENAGLEGSVIVERLKNAEPGVGFNAATNEWVNMLEAGIVDPTKVTRSALQHAASVA
                     AMFLTTEAVVASIPEKNNDQPNMGGMPGMM"
     gene            complement(5125. .5409)
                     /gene="groES"
     CDS             complement(5125. .5409)
                     /gene="groES"
                     /note="ORFID:MW1954"
                     /codon_start=1
                     /transl_table=11
                     /product="GroES protein"
                     /protein_id="BAB95819.1"
                     /db_xref="GI:21205125"
                     /translation="MLKPIGNRVIIEKKEQEQTTKSGIVLTDSAKEKSNEGVIVAVGT
                     GRLLNDGTRVTPEVKEGDRVVFQQYAGTEVKRDNETYLVLNEEDILAVIE"
     gene            5584. .6327
                     /gene="MW1955"
     CDS             5584. .6327
                     /gene="MW1955"
                     /note="ORFID:MW1955"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB95820.1"
                     /db_xref="GI:21205126"
                     /translation="MTRLWASLLTVIIYILSQFLPLLIVKKLPFVQYSGIELTKAVIY
                     IQLVLFLIAATTIILINLKIKNPTKLELEVKEPKKYIIPWALLGFALVMIYQMVVSIV
                     LTQIYGGQQVSPNTEKLIIIARKIPIFIFFVSIIGPLLEEYVFRKVIFGELFNAIKGN
                     RIVAFIIATTVSSLIFALAHNDFKFIPVYFGMGVIFSLAYVWTKRLAVPIIIHMLQNG
                     FVVIFQLLNPEALKKATEQANFIYHIFIP"
     gene            complement(6352. .7611)
                     /gene="MW1956"
     CDS             complement(6352. .7611)
                     /gene="MW1956"
                     /note="ORFID:MW1956
                     hypothetical protein, similar to SdrH"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB95821.1"
                     /db_xref="GI:21205127"
                     /translation="MSYHWFKKMLLSTSILILSSSSLGLATHTVEAKDNLNGEKPTTN
                     LNHNITSPSVNSEMNNNETGTPHESNQTGNEGTGSNSRDANPDSNNVKPDSNNQNPST
                     DSKPDPNNQNPSPNPKPDPDNPKPKPDPKPDPDKPKPNPDPKPDPDNPKPNPDPKPDP
                     DKPKPNPDPKPDPDKPKPNPNPKPDPNKPNPNPSPDPDQPGDSNHSGGSKNGGTWNPN
                     ASDGSNQGQWQPNGNQGNSQNPTGNDFVSQRFLALANGAYKYNPYILNQINKLGKDYG
                     EVTDEDIYNIIRKQNFSGNAYLNGLQQQSNYFRFQYFNPLKSERYYRNLDEQVLALIT
                     GEIGSMPDLKKPEDKPDSKQRSFEPHEKDDFTVVKKQEDNKKSASTAYSKSWLAIVCS
                     MMVVFSIMLFLFVKRNKKKNKNESQRR"
     gene            7808. .8434
                     /gene="MW1957"
     CDS             7808. .8434
                     /gene="MW1957"
                     /note="ORFID:MW1957"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB95822.1"
                     /db_xref="GI:21205128"
                     /translation="MGLFKKDKKAMTFDNAMEERRSIYNLKDSISISDDELESMIAHA
                     VKHVPSSFNSQSTRIVLLLNDNNNKFWDNTKAILKEVMGENRDFEPTEQKIDNFKHSY
                     GTILFYEDQDVVSGLQEQMPNYYDNFAIWSTQTNAMHQFAIWTALATKGIGASLQHYN
                     PLVDEMTSNEFNIPKSWKLIAQMPFGDIREAAGEKTFNPVEDRFVIKK"
     gene            8795. .9580
                     /gene="MW1958"
     CDS             8795. .9580
                     /gene="MW1958"
                     /note="ORFID:MW1958"

  Query Match             56.4%;  Score 569;  DB 1;  Length 304050;
  Best Local Similarity   73.3%;  Pred. No. 8.6e-78;
  Matches  728;  Conservative    0;  Mismatches  265;  Indels    0;  Gaps    0;

Qy          7 GAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCA 66
              | | |  |||||||| |||||| ||| ||| || | || ||| |  |||  ||| ||  |
Db     171480 GGAACGATGAGAGGTCTAAAAACTTTTAGTATATTGGGATTAATAGTTGCCTTACTTTTA 171421

Qy         67 ACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGA 126
                |||||| ||||| ||||    |      |    ||   ||   ||| | ||||||
Db     171420 GTTGCAGCTTGTGGTAATACGGATAATTCAAGTAAAAAAGAATCATCAACTAAAGATACT 171361

Qy        127 GTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTT 186
               ||    | ||  | ||| | |||||    |||||||||||| |  | |||||| | |||
Db     171360 ATTTCGGTAAAAGATGAAAATGGTACAGTAAAAGTACCTAAAGATGCAAAACGTATCGTT 171301

Qy        187 GTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATA 246
              ||  | ||||| ||||||| ||||| ||||  || ||||| ||||| || ||||| ||
Db     171300 GTATTAGAGTACTCATTTGCAGATGCATTAGCAGCATTAGACGTTAATCCAGTTGGTATT 171241

Qy        247 GCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATAC 306
              || ||||||   || ||||| ||||| ||||||||| | ||||| ||||||||  | ||
Db     171240 GCTGATGATGGTAAGAAAAAACGTATCATTAAACCAGTTAGAGAAAAAATTGGGGATTAT 171181

Qy        307 ACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGAT 366
              ||||||||||| |||||||| || || |||||||||||||| |||||| | ||||| |||
Db     171180 ACTTCTGTAGGTACACGTAAACAGCCAAACTTAGAAGAAATTAGTAAATTAAAACCGGAT 171121

Qy        367 TTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCT 426
              |||||||| |||||||  | |||||| ||||||||| ||||||| ||||| ||||||||
Db     171120 TTAATTATCGCTGATAGCAGTAGACATAAAGGTATTAATAAAGAATTAAACAAAATTGCA 171061

Qy        427 CCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAA 486
              || ||  |   | | || ||||| |||||||| || ||  |||| ||| || | || |||
Db     171060 CCAACATTATCATTAAAGAGTTTTGATGGAGACTACAAACAAAACATTAATTCGTTCAAA 171001

Qy        487 ACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGAT 546
              |||||| | |||||||||  ||||||| |||||||  |||||||  | |  || || |||
Db     171000 ACAATTGCTAAAGCTTTAAATAAAGAAAAAGAAGGCGAAAAACGTCTTGCTGAGCATGAT 170941
```

```
Qy        547 AAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCT 606
              ||   |||  |  | ||||||| | ||||| |   | |||| ||||||||| ||  | ||
Db     170940 AAATTAATCAATAAGTATAAAGATGAAATTAAATTTGATAGAAATCAAAAAGTGCTTCCA 170881

Qy        607 GCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTC 666
              ||||||| ||||||| | ||||||  | || ||||||| | | || ||||||||| |||||
Db     170880 GCAGTAGTTGCTAAAGCTGGTTTATTAGCACATCCAAACTATTCATATGTTGGACAATTT 170821

Qy        667 CTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAG 726
               |||   ||||||| ||||| | ||||||||||||| ||||| || ||||||||||||||
Db     170820 TTAAACGAACTAGGATTTAAAAATGCATTAAGTGACGATGTAACAAAAGGTTTAAGTAAA 170761

Qy        727 TATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAG 786
              ||| | ||||||||||||||||||| |  ||||||||  |||| || |  | ||||| |||
Db     170760 TATTTGAAAGGACCTTACTTACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAG 170701

Qy        787 CGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAA 846
              |||||| |||| |||||  |  | || |     |     ||  | | ||  || || ||
Db     170700 CGTATGATCATTATGACTGATCACGCTAAAAAAGATTCTGCTGAATTCAAGAAATTACAA 170641

Qy        847 AAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGAC 906
               ||||| |    |||||| ||||| || || || |||||| | || || |||||| | |||
Db     170640 GAAGATGCAACATGGAAAAAATTGAATGCAGTTAAAAATAATCGCGTGGATATTGTTGAC 170581

Qy        907 CGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTT 966
              |||||  | ||||||||||| ||||| |||||||||||| ||||||||||| |||||||||
Db     170580 CGTGATGTTTGGGCAAGATCTCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTT 170521

Qy        967 GTTGAATTATCTAAGAAAGATAGTAAAAAAGAT 999
              ||||||||||| || |||||    ||  ||| |
Db     170520 GTTGAATTATCAAAAAAAGAACAAAAGTAAGGT 170488


RESULT 4
AP003136/c
LOCUS       AP003136              307750 bp    DNA     linear   BCT 11-JAN-2003
DEFINITION  Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
            genome, section 8/10.
ACCESSION   AP003136 BA000018
VERSION     AP003136.2  GI:14349228
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus N315
  ORGANISM  Staphylococcus aureus subsp. aureus N315
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
            Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
            Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
            Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
            Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
            Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
            Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
  TITLE     Whole genome sequencing of meticillin-resistant Staphylococcus
            aureus
  JOURNAL   Lancet 357 (9264), 1225-1240 (2001)
  MEDLINE   21311952
   PUBMED   11418146
REFERENCE   2  (bases 1 to 307750)
  AUTHORS   Director-General, Biotechnology Center, Aoki,K., Oguchi,A.,
            Hosoyama,A., Nagai,Y., Kuroda,M., Hiramatsu,K. and Kikuchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JAN-2001) Director-General, Biotechnology Center,
            National Institute of Technology and Evaluation, Biotechnology
            Center; 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
            (E-mail:bio@nite.go.jp, URL:http://www.bio.nite.go.jp/,
            Tel:81-3-3481-1933, Fax:81-3-3481-8424)
COMMENT     On Jun 12, 2001 this sequence version replaced gi:13701842.
FEATURES             Location/Qualifiers
     source          1. .307750
                     /organism="Staphylococcus aureus subsp. aureus N315"
                     /mol_type="genomic DNA"
                     /strain="N315"
                     /sub_species="aureus N315"
                     /db_xref="taxon:158879"
     gene            297. .1907
                     /gene="SA1853"
     CDS             297. .1907
                     /gene="SA1853"
                     /note="ORFID:SA1853
                     hypothetical protein, similar to DNA mismatch repair
                     protein MutS"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB43135.1"
                     /db_xref="GI:13701843"
                     /translation="MSTNQTFLIFVIAIILLTSVIGIVGRYMSRQRLLKSMETLWQTI
                     SPLEAFIRPNSHFDYEYKLYKEKFESHSLVDDKTWSDLNMNAIFHKMNYNLTAIGEMK
                     LYACLRGMLSITNKSLLSLFNDNAEFRKNVTYHLALIGKTVYPTFPDQITPVKRQNIL
                     FLCPFLPVISFAVIFINSQVGILLFLMSCLFNIILSATLKRTYEDDLKSIFYASNVLK
                     QGYTISKIKHAPQPEVNFKQFRTARHLTSVLAEVNDEDIGAMVIKLVKLIFMLDYVLF
                     HSIQKSYTTHMNELKNCFDYIAELDNHYALAMYRRTLECYTEPQIDDTNDGIVFSELT
                     HPLIADAVANDFSLSQNILLTGSNASGKSTFMKSIAINIILASAIQTVTASKFVYQPG
                     IVFTSMANADDVLSGDSYFMAELKSIKRIVEIPDNQKIYCFIDEIFKGTNTTERIAAS
                     ESVLSFLHEKSNFRVIAATHDIELAELLKQRYENYHFNEVIENNNIHFDYKIKPGKAN
                     TRNAIELLKITSFPAKIYERAKDNVSNG"
     gene            complement(2222. .3247)
                     /gene="SA1854"
     CDS             complement(2222. .3247)
                     /gene="SA1854"
                     /note="ORFID:SA1854
                     hypothetical protein, similar to O-sialoglycoprotein
                     endopeptidase"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB43136.1"
                     /db_xref="GI:13701844"
                     /translation="MTKDILILAVETSCDETSVSVIKNGRDILSNTVLSQIESHKRFG
                     GVVPEVASRHHVEGITTTINEALVDADVSMEDIDAIAVTEGPGLIGALLIGVNAAKAL
                     AFAYDKPLIPVHHIAGHIYANHIEEPLTFPLIALIVSGGHTELVYMKDHLSFEVIGET
                     RDDAVGEAYDKVARTIGLNYPGGPQVDRLAAEGEDTYSFPRVWLDKDSYDFSFSGLKS
                     AVINQLHNQRQKNIPIIEANVATSFQNSVVEVLTFKAIQACKEYSVQRLIVAGGVASN
                     KGLRQSLADQCKVNDIQLTIPSPKLCTDNAAMIGVAGHSLYQQGRFADLALNGHSNID
                     LEEYSAE"
     gene            complement(3240. .3704)
                     /gene="SA1855"
     CDS             complement(3240. .3704)
                     /gene="SA1855"
                     /note="ORFID:SA1855
                     hypothetical protein, similar to ribosomal-protein-alanine
                     N-acetyltransfer"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB43137.1"
                     /db_xref="GI:13701845"
                     /translation="MDQQSKEQLNIREMTKEDVPQVFDLERRSFNDSSWTIDAFYHEI
                     EQNNFAKYFVLEFEQQIIGYLGLWIVIDQAQITTVAIDDNYRGYGLGQMLLKYGINYA
                     SHTCDVMSLEVRVNNKVAQHVYENLGFQYGGKRKNYYGEGEDAMVMWVNLND"
     gene            complement(3677. .4366)
                     /gene="SA1856"
     CDS             complement(3677. .4366)
                     /gene="SA1856"
                     /note="ORFID:SA1856
                     hypothetical protein, similar to glycoprotein
                     endopeptidase"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB43138.1"
                     /db_xref="GI:13701846"
                     /translation="MNKLRRQLSMNSLLIDTSNQPLSVALMQNDEVLAEITTDSKQNH
                     SVQLMPAISQLFEQSKIAKQQLDAIIVAEGPGSYTGLRIGVTVAKTLAYALDVKLYGV
                     SSLKALAATIDHTDKLLVPVFDARRQAVYTGIFQWQNGQLVTILEDQYMSIEALRAFL
                     EDLNQPFIYIGKDTIKLQDELQGEVIAQLPNASVMYQLIDKPTDVHTFTPKYHKLAEA
                     ERNWINSQKNN"
     gene            complement(4320. .4781)
                     /gene="SA1857"
     CDS             complement(4320. .4781)
                     /gene="SA1857"
                     /note="ORFID:SA1857"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB43139.1"
                     /db_xref="GI:13701847"
                     /translation="MIKINNLDEMNQFAMFLVEQLKSGDLILLNGDLGAGKTTLTQFI
                     GKALGVRRTINSPTFNIIKSYRGKNLKLHHMDCYRLEDSDEDLGFDEFFEDQAITVIE
                     WSQFIKDLLPATHLSINISTISENTRQIELFAQGEHYEQIKEAIIHEFAAH"
     gene            5292. .6980
                     /gene="ilvD"
     CDS             5292. .6980
                     /gene="ilvD"
                     /note="ORFID:SA1858"
                     /codon_start=1
                     /transl_table=11
                     /product="dihydroxy-acid dehydratase"
                     /protein_id="BAB43140.1"
                     /db_xref="GI:13701848"
                     /translation="MRSDMIKKGDHQAPARSLLHATGALKSPTDMNKPFVAICNSYID
                     IVPGHVHLRELADIAKEAIREAGAIPFEFNTIGVDDGIAMGHIGMRYSLPSREIIADA
                     AETVINAHWFDGVFYIPNCDKITPGMILAAMRTNVPAIFCSGGPMKAGLSAHGKALTL
                     SSMFEAVGAFKEGSISKEEFLDMEQNACPTCGSCAGMFTANSMNCLMEVLGLALPYNG
                     TALAVSDQRREMIRQAAFKLVENIKNDLKPRDIVTREAIDDAFALDMAMGGSTNTVLH
                     TLAIANEAGIDYDLERINAIAKRTPYLSKIAPSSSYSMHDVHEAGGVPAIINELMKKD
                     GTLHPDRITVTGKTLRENNEGKEIKNFDVIHPLDAPYDAQGGLSILFGNIAPKGAVIK
                     VGGVDPSIKTFTGKAICFNSHDEAVEAIDNRTVRAGHVVVIRYEGPKGGPGMPEMLAP
                     TSSIVGRGLGKDVALITDGRFSGATRGIAVGHISPEAASGGPIALIEDGDEITIDLTN
                     RTLNVNQPEDVLARRRESLTPFKAKVKTGYLARYTALVTSANTGGVMQVPENLI"
     gene            7008. .8777
                     /gene="ilvB"
     CDS             7008. .8777
                     /gene="ilvB"
                     /note="ORFID:SA1859"
                     /codon_start=1
                     /transl_table=11
                     /product="acetolactate synthase large subunit"
                     /protein_id="BAB43141.1"
                     /db_xref="GI:13701849"
                     /translation="MSKTQHEVNQNIDPLKMAESLELEQLNEKTLNDMRSGSEVLVEA
                     LLKENVDYLFGYPGGAVLPLYDTFYDGKIKHILARHEQGAVHAAEGYARVSGKTGVVV
                     VTSGPGATNVMTGITDAHCDSLPLVVFTGQVATPGIGKDAFQEADILSMTSPITKQNY
                     QVKRVEDIPKIVHEAFHVANSGRKGPVVIDFPKDMGVLATNVDLCDEINIPGYEVVTE
                     PENKDIDTFISLLKEAKKPVVLAGAGINQSKSNQLLTQFVNKHQIPTVTTLLGLGAVP
                     YEDTLFLGMGGMHGSYASNMALTECDLLINLGSRFDDRLASKPDAFAPNAKIVHVDID
                     PSEINKVIHVDLGIIADCKRFLECLNDKNVETIEHSDWVKHCQNNKQKHPFKLGEEDQ
                     VFCKPQQTIEYIGKITNGEAIVTTDVGQHQMWAAQFYPFKNHGQWVTSGGLGTMGFGI
                     PSSIGAKLANPDKTVVCFVGDGGFQMTNQEMALLPEYGLDVKIVLINNGTLGMVKQWQ
                     DKFFNQRFSHSVFNGQPDFMKMAEAYGVKGFLIDKPEQLEEQLDAAFAYQGPVLIEVR
                     ISPTEAVTPMVPSGKSNHEMEGL"
     gene            8777. .9031
                     /gene="SA1860"
     CDS             8777. .9031
                     /gene="SA1860"
                     /note="ORFID:SA1860
                     hypothetical protein, similar to acetolactate synthase
                     small subunit"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB43142.1"
                     /db_xref="GI:13701850"
                     /translation="MTRILKLQVADQVSTLNRITSAFVRLQYNIDTLHVTHSEQPGIS
                     NMEIQVDIQDDTSLHILIKKLKQQINVLTVECYDLVDNEA"
     gene            9168. .10172
                     /gene="ilvC"
     CDS             9168. .10172
                     /gene="ilvC"
                     /note="ORFID:SA1861"
                     /codon_start=1
                     /transl_table=11
                     /product="alpha-keto-beta-hydroxylacil reductoisomerase"
                     /protein_id="BAB43143.1"
                     /db_xref="GI:13701851"
                     /translation="MTTVYYDQDVKTDALQGKKIAVVGYGSQGHAHAQNLKDNGYDVV
                     IGIRPGRSFDKAKEDGFDVFPVAEAVKQADVIMVLLPDEIQGDVYKNEIEPNLEKHNA
                     LAFAHGFNIHFGVIQPPADVDVFLVAPKGPGHLVRRTFVEGSAVPSLFGIQQDASGQA
                     RNIALSYAKGIGATRAGVIETTFKEETETDLFGEQAVLCGGVSKLIQSGFETLVEAGY
                     QPELAYFEVLHEMKLIVDLMYEGGMENVRYSISNTAEFGDYVSGPRVITPDVKENMKA
                     VLTDIQNGNFSNRFIEDNKNGFKEFYKLREEQHGHQIEKVGRELREMMPFIKSKSIEK
                     "
     gene            10202. .11731
                     /gene="leuA"
     CDS             10202. .11731
                     /gene="leuA"
                     /note="ORFID:SA1862"
                     /codon_start=1
                     /transl_table=11
                     /product="2-isopropylmalate synthase"

  Query Match             56.4%;  Score 569;  DB 1;  Length 307750;
  Best Local Similarity   73.3%;  Pred. No. 8.6e-78;
  Matches  728;  Conservative    0;  Mismatches  265;  Indels    0;  Gaps    0;

Qy          7 GAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCA 66
              | | |  |||||||| ||||||| ||| ||| || | || |  |||   |||  ||| ||  |
Db     156657 GGAACGATGAGAGGTCTAAAAACTTTTAGTATATTGGGATTAATAGTTGCCTTACTTTTA 156598

Qy         67 ACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGA 126
                |||||| ||||| |||| |    |      |    ||   ||   ||| | ||||||   
Db     156597 GTTGCAGCTTGTGGTAATACGGATAATTCAAGTAAAAAAGAATCATCAACTAAAGATACT 156538

Qy        127 GTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTT 186
               ||    | ||  | ||| | |||||    |||||||||||| |  | |||||| | |||
Db     156537 ATTTCGGTAAAAGATGAAAATGGTACAGTAAAAGTACCTAAAGATGCAAAACGTATCGTT 156478

Qy        187 GTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATA 246
              ||  | |||||| ||||||| ||||| ||||  || |||||| |||||||| ||||| || 
Db     156477 GTATTAGAGTACTCATTTGCAGATGCATTAGCAGCATTAGACGTTAAACCAGTTGGTATT 156418

Qy        247 GCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATAC 306
              || ||||||   || ||||| ||||| |||||||||| | ||||| ||||||||  | || 
Db     156417 GCTGATGATGGTAAGAAAAAACGTATCATTAAACCAGTTAGAGAAAAAATTGGGGATTAT 156358

Qy        307 ACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGAT 366
              ||||||||||| |||||||| || || |||||||| |||| |||||| | ||||| |||
Db     156357 ACTTCTGTAGGTACACGTAAACAGCCAAACTTAGAGGAAATTAGTAAATTAAAACCGGAT 156298

Qy        367 TTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCT 426
              |||||||| ||||||||  | |||||| |||||||| |||||||| ||||| |||||||| 
Db     156297 TTAATTATCGCTGATAGCAGTAGACATAAAGGTATTAATAAAGAATTAAACAAAATTGCA 156238

Qy        427 CCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAA 486
              || ||  |   | | || ||||| |||||||| || ||  |||||||| || | | || |||
Db     156237 CCAACATTATCATTAAAGAGTTTTGATGGAGACTACAAACAAAATATTAATTCGTTCAAA 156178

Qy        487 ACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGAT 546
              |||||| | |||||||||  ||||||| |||||||  |||||||  | |  || || |||
Db     156177 ACAATTGCTAAAGCTTTAAATAAAGAAAAAGAAGGCGAAAAACGTCTTGCTGAGCATGAT 156118

Qy        547 AAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCT 606
              ||   |||  |  | ||||||| | ||||| |   | ||||  ||||||||| ||  | || 
Db     156117 AAATTAATCAATAAGTATAAAGATGAAATTAAATTTGATAGAAATCAAAAAGTGCTTCCA 156058

Qy        607 GCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTC 666
              ||||||| ||||||| | |||||  | || ||||||| | | || |||||||||| ||||| 
Db     156057 GCAGTAGTTGCTAAAGCTGGTTTATTAGCACATCCAAACTATTCATATGTTGGACAATTT 155998

Qy        667 CTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAG 726
               |||   ||||||| ||||||| | ||||||||||| ||||| || |||||||||||||| 
Db     155997 TTAAACGAACTAGGATTTAAAAATGCATTAAGTGACGATGTAACAAAAGGTTTAAGTAAA 155938

Qy        727 TATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAG 786
              ||| | |||||||||||||||||| |  ||||||| | |||| || |  | ||||| |||
Db     155937 TATTTGAAAGGACCTTACTTACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAG 155878
```

```
Qy         787 CGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAA 846
               |||||| |||| |||||| |  | || |      |     ||  | | ||  |  || ||
Db      155877 CGTATGATCATTATGACAGATCATGCTAAAAAAGATTCTGCTGAATTCAAGAAGTTACAA 155818

Qy         847 AAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGAC 906
                ||||| |   |||||| || || || || || |||||| | || || |||||| | |||
Db      155817 GAAGATGCAACATGGAAAAAGTTGAATGCAGTTAAAAATAATCGCGTGGATATTGTTGAC 155758

Qy         907 CGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTT 966
               |||||  | ||||||||||| ||||| ||||||||||| ||||||||||| |||||||||
Db      155757 CGTGATGTTTGGGCAAGATCTCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTT 155698

Qy         967 GTTGAATTATCTAAGAAAGATAGTAAAAAAGAT 999
               ||||||||||| || |||||    ||  ||| |
Db      155697 GTTGAATTATCAAAAAAAGAACAAAAGTAAGGT 155665


RESULT 5
AP003364/c
LOCUS       AP003364               347650 bp    DNA     linear   BCT 17-APR-2004
DEFINITION  Staphylococcus aureus subsp. aureus Mu50 DNA, complete genome,
            section 7/9.
ACCESSION   AP003364 BA000017
VERSION     AP003364.3  GI:46395538
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus Mu50
  ORGANISM  Staphylococcus aureus subsp. aureus Mu50
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
            Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
            Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
            Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
            Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
            Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
            Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
  TITLE     Whole genome sequencing of meticillin-resistant Staphylococcus
            aureus
  JOURNAL   Lancet 357 (9264), 1225-1240 (2001)
  MEDLINE   21311952
   PUBMED   11418146
REFERENCE   2
  AUTHORS   Ohta,T., Hirakawa,H., Morikawa,K., Maruyama,A., Inose,Y.,
            Yamashita,A., Oshima,K., Kuroda,M., Hattori,M., Hiramatsu,K.,
            Kuhara,S. and Hayashi,H.
  TITLE     Nucleotide Substitutions in Staphylococcus aureus Strains, Mu50,
            Mu3, and N315
  JOURNAL   DNA RESEARCH 11, 51-56 (2004)
REFERENCE   3  (bases 1 to 347650)
  AUTHORS   Hattori,M., Yamashita,A., Oshima,K. and Shiba,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Masahira Hattori, Kitasato Institute for
            Life Sciences, Kitasato University; Kitasato 1-15-1, Sagamihara,
            Kanagawa 228-8555, Japan
            (E-mail:hattori@genome.ls.kitasato-u.ac.jp, Tel:81-42-778-8194,
            Fax:81-42-778-8193)
COMMENT     On Apr 16, 2004 this sequence version replaced gi:14247707.
            This work was done in collaboration with Toshiko Ohta, Mutsumi
            Kanamori, Hideo Hayashi (University of Tsukuba), Hideki Hirakawa,
            Satoru Kuhara (Kyushu University), Ikuo Uchiyama (Okazaki National
            Research Institutes), Susumu Goto, Minoru Kanehisa (Kyoto
            University), Naotake Ogasawara (Nara Institute of Science and
            Technology) and supported by the Research for the Future Program of
            the Japan Society for the Promotion of Science.
FEATURES             Location/Qualifiers
     source          1. .347650
                     /organism="Staphylococcus aureus subsp. aureus Mu50"
                     /mol_type="genomic DNA"
                     /strain="Mu50"
                     /sub_species="aureus Mu50"
                     /db_xref="taxon:158878"
     gene            157. .1443
```

```
                     /locus_tag="SAV1936"
     CDS             157. .1443
                     /locus_tag="SAV1936"
                     /codon_start=1
                     /transl_table=11
                     /product="similar to aspartate transaminase protein"
                     /protein_id="BAB58098.1"
                     /db_xref="GI:14247709"
                     /translation="MNPLAQSLNEQLQQSNATAFTMLSDLGQNMFYPKGILSQSAEAK
                     STTYNATIGMATNKDGKMFASSLDAMFNDLTPDEIFPYAPPQGIEELRDLWQQKMLRD
                     NPELSIDNMSLPIVTNALTHGLSLVGDLFVNQGDTILLPEHNWGNYKLVFNTRNGANL
                     QTYPIFDKDGHYTTDSLVEALQSYNKDKVIMILNYPNNPTGYTPTHKEVTTIVDAIKA
                     LADKGTKVIAVVDDAYYGLFYEDVYTQSLFTALSNLNSNAILPVRLDGATKEFFAWGF
                     RVGFMTFGTSDQTTKEVLEAKVKGLIRSNISSGPLPTQSAVKHVLKNNKQFDKEIEQN
                     IQTLKERYEVTKEVVYADQYHSHWQAYDFNSGYFMAIKVHDVDPEELRKHLIEKYSIG
                     VIALNATDIRIAFSCVEKEDIPHVFDSIAKAIDDLR"
     gene            complement(1722. .2042)
                     /locus_tag="SAV1937"
     CDS             complement(1722. .2042)
                     /locus_tag="SAV1937"
                     /note="Bacteriophage phi Sa 3mu"
                     /codon_start=1
                     /transl_table=11
                     /product="truncated map-w protein"
                     /protein_id="BAB58099.1"
                     /db_xref="GI:14247710"
                     /translation="MKKKINKSNNVKFPVTINKFENIVSNEFVFYNASKITINDLSIK
                     LKSAMANDQGITKHDIGLAERAVYKVYFKNGSSKYVDLKTEYKDERVFKATDIKKVDI
                     ELKF"
     gene            complement(2027. .3457)
                     /locus_tag="SAV1938"
     CDS             complement(2027. .3457)
                     /locus_tag="SAV1938"
                     /note="Bacteriophage phi Sa 3mu"
                     /codon_start=1
                     /transl_table=11
                     /product="truncated map-w protein"
                     /protein_id="BAB58100.1"
                     /db_xref="GI:14247711"
                     /translation="MKFKSLITTTLALGVIASTGANFNTNEASAAAKPLDKSSSTLHH
                     GHSNIQIPYTITVNGTSQNILSSLTFNKNQNISYKDIENKVKSVLYFNRGISDIDLRL
                     SKQAEYTVHFKNGTKRVIDLKSGIYTADLINTSDIKAISVNVDTKKQPKDKAKANVQV
                     PYTITVNGTSQNILSNLTFNKNQNISYKDLEGKVKSVLESNRGITDVDLRLSKQAKYT
                     VNFKNGTKKVIDLKSGIYTANLINSSDIKSININVDTKKHIENKAKRNYQVPYSINLN
                     GTSTNILSNLSFSNKPWTNYKNLTSQIKSVLKHDRGISEQDLKYAKKAYYTVYFKNGG
                     KRILQLNSKNYTANLVHAKDVKRIEITVKTGTKAKADRYVPYTIAVNGTSTPILSKLK
                     ISNKQLISYKYLNDKVKSVLKSERGISDLDLKFAKQAKYTVYFKNGKKQVVNLKSDIF
                     TPNLFSAKDIKKIDIDVKQYTKSKKK"
     gene            3803. .4042
                     /gene="truncated-hlb"
                     /locus_tag="SAV1939"
     CDS             3803. .4042
                     /gene="truncated-hlb"
                     /locus_tag="SAV1939"
                     /note="Bacteriophage phi Sa 3mu"
                     /codon_start=1
                     /transl_table=11
                     /product="truncated beta-hemolysin"
                     /protein_id="BAB58101.2"
                     /db_xref="GI:46395539"
                     /translation="MHLFCCKLYKRSDNDGKKTKSNSLKKVATLALANLLLVGALTDN
                     SAKAESKKDDTDLKLVSHNVYMLSTVLYPNWRLLT"
     gene            4414. .4593
                     /locus_tag="SAV1940"
     CDS             4414. .4593
                     /locus_tag="SAV1940"
                     /note="Bacteriophage phi Sa 3mu"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB58102.1"
                     /db_xref="GI:14247713"
                     /translation="MKFKTTKECKSNNIFKRSQEINNSESEKGCLWGISMLILLFLLI
                     LFGITACSSSIHFIN"
     gene            4617. .4925
                     /locus_tag="SAV1941"
     CDS             4617. .4925
                     /locus_tag="SAV1941"
                     /note="Bacteriophage phi Sa 3mu"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB58103.2"
                     /db_xref="GI:46395540"
                     /translation="MKNHTNIINILLVIVNSLTHFLTLNTSFFNNSASDFCFIIGAIF
                     FLIGIFVAIYGMKRATYWLNLLILFTNIFYFLHFCVLLLLKYIGFKLFIYEGCVLLFT
                     "
     gene            complement(5216. .5566)
                     /locus_tag="SAV1942"
     CDS             complement(5216. .5566)
                     /locus_tag="SAV1942"
                     /note="Bacteriophage phi Sa 3mu"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB58104.1"
                     /db_xref="GI:14247715"
                     /translation="MKIRKSILAGTLAIVLASPLVTNLDKNEAQASTSLPTSNEYQNE
                     KLANELKSLLDELNVNELATGSLNTYYKRTIKISGQKAMYALKSKDFKKMSEAKYQLQ
                     KIYNEIDEALKSKY"
     gene            complement(6077. .6373)
                     /gene="truncated-lytA"
                     /locus_tag="SAV1943"
     CDS             complement(6077. .6373)
                     /gene="truncated-lytA"
                     /locus_tag="SAV1943"
                     /note="Bacteriophage phi Sa 3mu"
                     /codon_start=1
                     /transl_table=11
                     /product="truncated amidase"
                     /protein_id="BAB58105.2"
                     /db_xref="GI:46395541"
                     /translation="MPAGYTLDKNNVPYKKETGYYTVANVKGNNVRDGYSTNSRITGV
                     LPNNATIKYDGAYCINGYRWITYIANNGQRCYIATGEVDKAGNRISSFGNFSAL"
     gene            6461. .6616
                     /gene="truncated-lytA"
                     /locus_tag="SAV1943.1"
     CDS             6461. .6616
                     /gene="truncated-lytA"
                     /locus_tag="SAV1943.1"
                     /note="Bacteriophage phi Sa 3mu"
                     /codon_start=1
                     /transl_table=11
                     /product="truncated amidase"
                     /protein_id="BAD16608.1"
                     /db_xref="GI:46395542"
                     /translation="MRCSDCNISMLLRYMYRVKGKKDINCRISVRLLGHLTGYIVHSY
                     YIHLIIT"
     gene            complement(7063. .7554)
                     /gene="sak"
                     /locus_tag="SAV1944"
     CDS             complement(7063. .7554)
                     /gene="sak"
                     /locus_tag="SAV1944"
                     /note="Bacteriophage phi Sa 3mu"
                     /codon_start=1
                     /transl_table=11
                     /product="staphylokinase precursor"
                     /protein_id="BAB58106.1"
                     /db_xref="GI:14247717"
                     /translation="MLKRSLLFLTVLLLLFSFSSITNEVSASSSFDKGKYKKGDDASY
                     FEPTGPYLMVNVTGVDGKGNELLSPHYVEFPIKPGTTLTKEKIEYYVEWALDATAYKE
                     FRVVELDPSAKIEVTYYDKNKKKEETKSFPITEKGFVVPDLSEHIKNPGFNLITKVVI
                     EKK"
     gene            complement(7769. .8500)
                     /locus_tag="SAV1945"
```

```
     CDS             complement(7769. .8500)

  Query Match             56.4%;  Score 569;  DB 1;  Length 347650;
  Best Local Similarity   73.3%;  Pred. No. 8.3e-78;
  Matches  728;  Conservative    0;  Mismatches  265;  Indels    0;  Gaps    0;

Qy          7 GAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCA 66
              | | |  |||||||| |||||| ||| ||| || | || ||| |  |||  ||| ||  |
Db     239356 GGAACGATGAGAGGTCTAAAAACTTTTAGTATATTGGGATTAATAGTTGCCTTACTTTTA 239297

Qy         67 ACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGA 126
                |||||| ||||| ||||    |      |    ||   ||   ||| | ||||||
Db     239296 GTTGCAGCTTGTGGTAATACGGATAATTCAAGTAAAAAAGAATCATCAACTAAAGATACT 239237

Qy        127 GTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTT 186
               ||    | ||  | ||| | |||||    |||||||||||| |  | |||||| | |||
Db     239236 ATTTCGGTAAAAGATGAAAATGGTACAGTAAAAGTACCTAAAGATGCAAAACGTATCGTT 239177

Qy        187 GTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATA 246
              ||  | ||||| |||||||  ||||| ||||  || ||||| |||||||| ||||| ||
Db     239176 GTATTAGAGTACTCATTTGCAGATGCATTAGCAGCATTAGACGTTAAACCAGTTGGTATT 239117

Qy        247 GCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATAC 306
              || ||||||   || ||||| ||||| ||||||||| | ||||| ||||||||  | ||
Db     239116 GCTGATGATGGTAAGAAAAAAACGTATCATTAAACCAGTTAGAGAAAAAATTGGGGATTAT 239057

Qy        307 ACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGAT 366
              ||||||||||| |||||||| || || |||||||| ||||| |||||| | ||||| |||
Db     239056 ACTTCTGTAGGTACACGTAAACAGCCAAACTTAGAGGAAATTAGTAAATTAAAACCGGAT 238997

Qy        367 TTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCT 426
              |||||||| |||||||  | |||||| ||||||||| ||||||| ||||| ||||||||
Db     238996 TTAATTATCGCTGATAGCAGTAGACATAAAGGTATTAATAAAGAATTAAACAAAATTGCA 238937

Qy        427 CCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAA 486
              || ||  |   | | || ||||| |||||||| || ||  |||||||| || | || |||
Db     238936 CCAACATTATCATTAAAGAGTTTTGATGGAGACTACAAACAAAATATTAATTCGTTCAAA 238877

Qy        487 ACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGAT 546
              |||||| | |||||||||  ||||||| |||||||  |||||||  | |  || || |||
Db     238876 ACAATTGCTAAAGCTTTAAATAAAGAAAAAGAAGGCGAAAAACGTCTTGCTGAGCATGAT 238817

Qy        547 AAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCT 606
              ||   |||  |  | |||||| | ||||| |   | |||| ||||||||| ||  | ||
Db     238816 AAATTAATCAATAAGTATAAAGATGAAATTAAATTTGATAGAAATCAAAAAGTGCTTCCA 238757

Qy        607 GCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTC 666
              ||||||| ||||||| | |||||  | || ||||||| | | || |||||||| |||||
Db     238756 GCAGTAGTTGCTAAAGCTGGTTTATTAGCACATCCAAACTATTCATATGTTGGACAATTT 238697

Qy        667 CTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAG 726
               |||   ||||||| |||||| | ||||||||||| ||||| || ||||||||||||||
Db     238696 TTAAACGAACTAGGATTTAAAAATGCATTAAGTGACGATGTAACAAAAGGTTTAAGTAAA 238637

Qy        727 TATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAG 786
              ||| | |||||||||||||||||| |  ||||||||  |||| || |  | ||||| |||
Db     238636 TATTTGAAAGGACCTTACTTACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAG 238577

Qy        787 CGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAA 846
              |||||| |||| |||||| |  | || |      |     ||  | | ||  |  || ||
Db     238576 CGTATGATCATTATGACAGATCATGCTAAAAAAGATTCTGCTGAATTCAAGAAGTTACAA 238517

Qy        847 AAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGAC 906
               ||||| |   |||||| || || || || |||||| | || || |||||| | |||
Db     238516 GAAGATGCAACATGGAAAAAGTTGAATGCAGTTAAAAATAATCGCGTGGATATTGTTGAC 238457

Qy        907 CGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTT 966
              |||||  | ||||||||||| ||||| ||||||||||| ||||||||||| |||||||||
Db     238456 CGTGATGTTTGGGCAAGATCTCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTT 238397

Qy        967 GTTGAATTATCTAAGAAAGATAGTAAAAAAGAT 999
              ||||||||||| || |||||    ||  ||| |
Db     238396 GTTGAATTATCAAAAAAAGAACAAAAGTAAGGT 238364
```

```
RESULT 6
BX571856_23/c
WPCOMMENT
  Sequence split into 29 fragments  LOCUS BX571856 Accession BX571856
     Fragment Name       Begin      End
     BX571856_00             1     110000
     BX571856_01        100001     210000
     BX571856_02        200001     310000
     BX571856_03        300001     410000
     BX571856_04        400001     510000
     BX571856_05        500001     610000
     BX571856_06        600001     710000
     BX571856_07        700001     810000
     BX571856_08        800001     910000
     BX571856_09        900001    1010000
     BX571856_10       1000001    1110000
     BX571856_11       1100001    1210000
     BX571856_12       1200001    1310000
     BX571856_13       1300001    1410000
     BX571856_14       1400001    1510000
     BX571856_15       1500001    1610000
     BX571856_16       1600001    1710000
     BX571856_17       1700001    1810000
     BX571856_18       1800001    1910000
     BX571856_19       1900001    2010000
     BX571856_20       2000001    2110000
     BX571856_21       2100001    2210000
     BX571856_22       2200001    2310000
     BX571856_23       2300001    2410000
     BX571856_24       2400001    2510000
     BX571856_25       2500001    2610000
     BX571856_26       2600001    2710000
     BX571856_27       2700001    2810000
     BX571856_28       2800001    2902619
Continuation (24 of 29) of BX571856 from base 2300001 (BX571856 Staphylococcus aureus su

  Query Match             56.3%;  Score 567.4;  DB 1;  Length 110000;
  Best Local Similarity   73.2%;  Pred. No. 2e-77;
  Matches  727;  Conservative    0;  Mismatches  266;  Indels    0;  Gaps    0;

Qy          7 GAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCA 66
              | | |  |||||||| |||||| ||| ||| || | || ||| |  |||  ||| ||  |
Db      50327 GGAACGATGAGAGGTCTAAAAACTTTTAGTATATTGGGATTAATAGTTGCCTTATTTTTA 50268

Qy         67 ACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGA 126
                |||||| ||||| ||||    |      |    ||   ||   ||| | ||||||
Db      50267 GTTGCAGCTTGTGGTAATACGGATAATTCAAGTAAAAAAGAATCATCAACTAAAGATACT 50208

Qy        127 GTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTT 186
               ||    | ||  | ||| | |||||    |||||||||||| |  | |||||| | |||
Db      50207 ATTTCGGTAAAAGATGAAAATGGTACAGTAAAAGTACCTAAAGATGCAAAACGTATCGTT 50148

Qy        187 GTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATA 246
              ||  | ||||| |||||||  ||||| ||||  || ||||| |||||||| ||||| ||
Db      50147 GTATTAGAGTACTCATTTGCAGATGCATTAGCAGCATTAGACGTTAAACCAGTTGGTATT 50088

Qy        247 GCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATAC 306
              || ||||||   || ||||| ||||| ||||||||| | ||||| |||||| |  | ||
Db      50087 GCTGATGATGGTAAGAAAAAAACGTATCATTAAACCAGTTAGAGAAAAAATTCGGGATTAT 50028

Qy        307 ACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGAT 366
              ||||||||||| |||||||| || || |||||||| ||||| |||||| | ||||| |||
Db      50027 ACTTCTGTAGGTACACGTAAACAGCCAAACTTAGAGGAAATTAGTAAATTAAAACCGGAT 49968

Qy        367 TTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCT 426
              |||||||| |||||||  | |||||| ||||||||| ||||||| ||||| ||||||||
Db      49967 TTAATTATCGCTGATAGCAGTAGACATAAAGGTATTAATAAAGAATTAAACAAAATTGCA 49908

Qy        427 CCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAA 486
              || ||  |   | | || ||||| ||||| || || ||  ||||||| |||||||| ||
Db      49907 CCAACATTATCATTAAAGAGTTTTGATGGTGACTACAAGCAAAATATCGATGCTTTCAAG 49848
```

```
Qy        487 ACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGAT 546
              |||||| | |||||||||| ||||||| |||||||  |||||||  | |  ||||| |||
Db      49847 ACAATTGCGAAAGCTTTAGATAAAGAAAAAGAAGGCGAAAAACGTCTTGCTGAACATGAT 49788

Qy        547 AAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCT 606
              ||   |||  |  | |||||| | ||||| |   | |||| ||||||||| ||  | ||
Db      49787 AAATTAATCAATAAGTATAAAGATGAAATTAAATTTGATAGAAATCAAAAAGTGCTTCCA 49728

Qy        607 GCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTC 666
              ||||| | ||| ||| | |||||  | || ||||||| | | || |||||||| |||||
Db      49727 GCAGTTGTTGCGAAAGCTGGTTTATTAGCACATCCAAACTATTCGTATGTTGGACAATTT 49668

Qy        667 CTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAG 726
               |||   |||| || |||||| | ||||||||||||||||| || ||||||||||||||
Db      49667 TTAAACGAACTTGGATTTAAAAATGCATTAAGTGATGATGTAACAAAAGGTTTAAGTAAA 49608

Qy        727 TATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAG 786
              ||  | |||||||||||||||||| |  | ||||||  |||| || |  | ||||||||
Db      49607 TACTTGAAAGGACCTTACTTACAATTAGATACTGAACATTTAGCTGACTTAAATCCTGAA 49548

Qy        787 CGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAA 846
              || ||| | || |||||| | || || |      |     ||  | | ||  |  || ||
Db      49547 CGCATGATTATTATGACAGATAATGCTAAAAAAGATTCTGCTGAATTCAAGAAGTTACAA 49488

Qy        847 AAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGAC 906
               ||||| |   |||||| ||||| || || || |||||| | |||||||||||| | |||
Db      49487 GAAGATGCAACATGGAAAAAATTGAATGCAGTTAAAAATAATCGTGTTGATATTGTTGAC 49428

Qy        907 CGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTT 966
              |||||  | ||||||||||| ||||| ||||||||||| ||||||||||| |||||||||
Db      49427 CGTGATGTTTGGGCAAGATCTCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTT 49368

Qy        967 GTTGAATTATCTAAGAAAGATAGTAAAAAAGAT 999
              ||||||||||| || |||||    ||  ||| |
Db      49367 GTTGAATTATCAAAAAAAGAACAAAAGTAAGGT 49335


RESULT 7
AX622368
LOCUS       AX622368                1044 bp    DNA     linear   PAT 20-FEB-2003
DEFINITION  Sequence 5331 from Patent WO02094868.
ACCESSION   AX622368
VERSION     AX622368.1  GI:28450503
KEYWORDS    .
SOURCE      Staphylococcus aureus
  ORGANISM  Staphylococcus aureus
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Masignani,V.C., Mora,M.C. and Scarselli,M.C.
  TITLE     Staphylococcus aureus proteins and nucleic acids
  JOURNAL   Patent: WO 02094868-A 5331 28-NOV-2002;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1. .1044
                     /organism="Staphylococcus aureus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1280"
ORIGIN

  Query Match             55.3%;  Score 557.4;  DB 6;  Length 1044;
  Best Local Similarity   73.5%;  Pred. No. 2.1e-75;
  Matches  711;  Conservative    0;  Mismatches  256;  Indels    0;  Gaps    0;

Qy         20 GTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTG 79
              | |  |||  ||| ||| || | || ||| |  |||  ||| ||  |  |||||| ||||
Db         71 GGTCTAAACCTTTTAGTATATTGGGATTAATAGTTGCCTTACTTTTAGTTGCAGCTTGTG 130

Qy         80 GAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGC 139
              | ||||    |      |    ||   ||   ||| | ||||||    ||    | ||
Db        131 GTAATACGGATAATTCAAGTAAAAAAGAATCATCAACTAAAGATACTATTTCGGTAAAAG 190

Qy        140 ACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATT 199
```

```
              | ||| | |||||    |||||||||||| |  | |||||| | |||||  | ||||| |
Db        191 ATGAAAATGGTACAGTAAAAGTACCTAAAGATGCAAAACGTATCGTTGTATTAGAGTACT 250

Qy        200 CATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACA 259
              ||||||  ||||| ||||  || ||||| |||||||| ||||| || || ||||||   |
Db        251 CATTTGCAGATGCATTAGCAGCATTAGACGTTAAACCAGTTGGTATTGCTGATGATGGTA 310

Qy        260 AAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAA 319
              | ||||| |||| |||||||||| | |||||| ||||||||  | || |||||||||||| |
Db        311 AGAAAAAACGTATCATTAAACCAGTTAGAGAAAAAAATTGGGGATTATACTTCTGTAGGTA 370

Qy        320 CACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTG 379
              ||||||| || || |||||||||||||||| |||||| | |||| |||||||||||| ||||
Db        371 CACGTAAACAGCCAAACTTAGAAGAAATTAGTAAATTAAAACCGGATTTAATTATCGCTG 430

Qy        380 ATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAAC 439
              |||  | |||||| |||||||||| ||||||| ||||| ||||||||| || ||  |   | 
Db        431 ATAGCAGTAGACATAAAGGTATTAATAAAGAATTAAACAAAATTGCACCAACATTATCAT 490

Qy        440 TGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAG 499
              | || ||||| |||||||| || ||  |||||||| || | || |||||||||| | ||||
Db        491 TAAAGAGTTTTGATGGAGACTACAAACAAAATATTAATTCGTTCAAAACAATTGCTAAAG 550

Qy        500 CTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAG 559
              |||||  ||||||| |||||||  |||||||  | |  || || |||||   |||  |  
Db        551 CTTTAAATAAAGAAAAAGAAGGCGAAAAACGTCTTGCTGAGCATGATAAATTAATCAATA 610

Qy        560 AATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTA 619
              | |||||| | ||||| |   | |||| ||||||||| ||  | || |||||| |||||
Db        611 AGTATAAAGATGAAATTAAATTTGATAGAAATCAAAAAGTGCTTCCAGCAGTAGTTGCTA 670

Qy        620 AATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAG 679
              || | |||||  | || ||||||| | | || |||||||| |||||  |||   ||||||
Db        671 AAGCTGGTTTATTAGCACATCCAAACTATTCATATGTTGGACAATTTTTAAACGAACTAG 730

Qy        680 GTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGAC 739
              | |||||| | ||||||||||| ||||| || ||||||||||||||| ||| | |||||||
Db        731 GATTTAAAAATGCATTAAGTGACGATGTAACAAAAGGTTTAAGTAAATATTTGAAAGGAC 790

Qy        740 CTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAA 799
              ||||||||||| |  ||||||||  |||| || |  | |||| |||||||||| |||| |
Db        791 CTTACTTACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGATCATTA 850

Qy        800 TGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTAT 859
              ||||| |  | || |       |     ||  | | ||  |  || || ||||| |   ||
Db        851 TGACAGATCATGCTAAAAAAGATTCTGCTGAATTCAAGAAGTTACAAGAAGATGCAACAT 910

Qy        860 GGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGG 919
              |||| || || || || || |||||| | || || ||||| | ||||||||  | ||||
Db        911 GGAAAAAGTTGAATGCAGTTAAAAATAATCGCGTGGATATTGTTGACCGTGATGTTTGGG 970

Qy        920 CAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTA 979
              ||||||| ||||| |||||||||||| ||||||||||| |||||||||||||||||||| |
Db        971 CAAGATCTCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTTGTTGAATTATCAA 1030

Qy        980 AGAAAGA 986
              | |||||
Db       1031 AAAAAGA 1037

RESULT 8
A64853/c
LOCUS       A64853                  2247 bp    DNA     linear   PAT 29-MAR-1999
DEFINITION  Sequence 10 from Patent WO9731114.
ACCESSION   A64853
VERSION     A64853.1  GI:4530844
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1
  AUTHORS   Burnham,M.K. and Hodgson,J.E.
  TITLE     POLYNUCLEOTIDES AND AMINOACID SEQUENCES FROM STAPHYLOCOCCUS AUREUS
  JOURNAL   Patent: WO 9731114-A 10 28-AUG-1997;
            SMITHKLINE BEECHAM PLC (GB)
FEATURES             Location/Qualifiers
     source          1. .2247
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             33.4%;  Score 336.2;  DB 6;  Length 2247;
  Best Local Similarity   70.6%;  Pred. No. 6.8e-42;
  Matches  580;  Conservative    0;  Mismatches  229;  Indels   13;  Gaps   10;

Qy          1 GGAGTGGAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTA 60
              ||||||| |    ||||||||| |||||| ||| ||| || | || ||| |  |||  |||
Db        831 GGAGTGGGACNGATGAGAGGTCTAAAAACTTTTAGTATATTGGGATTAATAGTTGCCTTA 772

Qy         61 ATTGCAACTGCAGCATGTGGAAATA---ATAGTTCAAGTAACTCAAGTAAAGAGTCATCA 117
               ||  |  |||||| ||||| ||||   ||| ||||||||||   ||| ||  |   |  |
Db        771 CTTTTAGTTGCAGCTTGTGGTAATACGGATAATTCAAGTAAAAGAAG-AATCATCAACTA 713

Qy        118 AAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAA 177
              ||      | ||    | ||  | ||| | |||||    ||||||||||||| |  |  ||
Db        712 AAGATACTATTTCGGGTAAAAGATGAAAATGGTACAGTAAAAGTACCTAAAGATGCACAA 653

Qy        178 CGTGTTGTTGTTCTTGAGTATTCATTTG--TTGATGCGTTAGTTGC-TTTAGATGTTAAA 234
              ||| | |||||  | ||||| |||||||    | ||| |||   || | |||  |||| |
Db        652 CGTATCGTTGTATTAGAGTACTCATTTGCAAGGTTGCATTANCAGCATATAGCCGTTANA 593

Qy        235 CC-TGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAA 293
              ||  ||||| || || ||||||   || ||||| ||||| | ||| ||| ||||||| ||
Db        592 CCAAGTTGGTATTGCTGATGATGGTAAGAAAAAACGTATCA-TAACCCAGTAAGAGA-AA 535

Qy        294 AATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAA 353
              ||||||     | ||||||||||| ||||||||| ||| |  ||||||  ||| || |||||
Db        534 AATTGGGGGTAATACTTCTGTAGGTACACGTAA-CAAGCNAACTTAAGAGANATTAGTAA 476

Qy        354 ACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTT 413
              | | || || |  |||||||||  ||||||  | |||||| || |||||| ||||||| ||
Db        475 ATTAAANCCGGGATTAATTATCNCTGATAGCAGTAGACATAAGGGTATTAATAAAGAATT 416

Qy        414 AAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATAT 473
              |    | ||||| || ||  |   | | || ||||| |||||||| || |   | |||||
Db        415 AGCCCACATTGCACC-ACATTATCATTAAAGAGTTTTGATGGAGACTACACCCACAATAT 357

Qy        474 TGATGCTTTTAA-AACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCT 532
              | || | || || ||||||| | |||||||||  ||||||  |||||||    |||||  
Db        356 TAATTCGTTCAACAACAATTGCTAAAGCTTTAAATAAAGANAAAGAAGGCGNNAAACGTC 297

Qy        533 TAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATC 592
              | |  || || |||||   |||  |  | |||||| | |||||| |   | |||| |||||
Db        296 TTGCTGAGCATGATAAATTAATCAATAAGTATAAAGATGAAATTAAATTTGATAGAAATC 237

Qy        593 AAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTT 652
              | || ||  | || |||||| ||||||| | |||||  | || ||||||  | | || |
Db        236 ACAAAGTGCTTCCAGCAGTAGTTGCTAAAGCTGGTTTATTAGCACATCCACACTATTCAT 177

Qy        653 ATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTA 712
              ||||||| |||||  |||   | ||||| |||||| | ||||||||||| ||||| ||  
Db        176 ATGTTGGACAATTTTTAAACGANCTAGGATTTAAAAATGCATTAAGTGACGATGTAACAC 117

Qy        713 AAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTC 772
              ||||||||||||| ||| | |||||||||||||||||||| |  ||||||||  |||| || 
Db        116 AAGGTTTAAGTAAATATTTGAAAGGACCTTACTTACAATTAGACACTGAACATTTAGCTG 57

Qy        773 AAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAA 814
              |  | ||||| ||||||||| |||| |||||||| |  | || |
Db         56 ATTTAAATCCAGAGCGTATGATCATTATGACAGATCATGCTA 15

RESULT 9
AP003132
LOCUS       AP003132              298050 bp    DNA     linear   BCT 11-JAN-2003
```

```
DEFINITION  Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
            genome, section 4/10.
ACCESSION   AP003132 BA000018
VERSION     AP003132.2  GI:14349174
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus N315
  ORGANISM  Staphylococcus aureus subsp. aureus N315
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
            Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
            Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
            Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
            Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
            Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
            Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
  TITLE     Whole genome sequencing of meticillin-resistant Staphylococcus
            aureus
  JOURNAL   Lancet 357 (9264), 1225-1240 (2001)
  MEDLINE   21311952
   PUBMED   11418146
REFERENCE   2  (bases 1 to 298050)
  AUTHORS   Director-General, Biotechnology Center, Aoki,K., Oguchi,A.,
            Hosoyama,A., Nagai,Y., Kuroda,M., Hiramatsu,K. and Kikuchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JAN-2001) Director-General, Biotechnology Center,
            National Institute of Technology and Evaluation, Biotechnology
            Center; 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
            (E-mail:bio@nite.go.jp, URL:http://www.bio.nite.go.jp/,
            Tel:81-3-3481-1933, Fax:81-3-3481-8424)
COMMENT     On Jun 12, 2001 this sequence version replaced gi:13700734.
FEATURES             Location/Qualifiers
     source          1. .298050
                     /organism="Staphylococcus aureus subsp. aureus N315"
                     /mol_type="genomic DNA"
                     /strain="N315"
                     /sub_species="aureus N315"
                     /db_xref="taxon:158879"
     gene            280. .432
                     /gene="SA0792"
     CDS             280. .432
                     /gene="SA0792"
                     /note="ORFID:SA0792"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB42031.1"
                     /db_xref="GI:13700735"
                     /translation="MKSKSKQPPNKYVEAFKPYLLTLLYLAIFITLYLIYGSGDTHNN
                     FIYNEF"
     gene            448. .1905
                     /gene="dltA"
     CDS             448. .1905
                     /gene="dltA"
                     /note="ORFID:SA0793"
                     /codon_start=1
                     /transl_table=11
                     /product="D-alanine-D-alanyl carrier protein ligase"
                     /protein_id="BAB42032.1"
                     /db_xref="GI:13700736"
                     /translation="MTDIINKLQAFADANPQSIAVRHTTDELTYQQLMDESSKLAHRL
                     QGSKKPMILFGHMSPYMIVGMIGAIKAGCGYVPVDTSIPEDRIKMIINKVQPEFVFNT
                     TDESFESLEGEVFTIEDIKTSQDPVIFDSQIKDNDTVYTIFTSGSTGEPKGVQIEYAS
                     LVQFTEWMLELNKSGNKQQWLNQAPFSFDLSVMAIYPCLASGGTLNLVDKNMINKPKL
                     LNEMLTATPINIWVSTPSFMEMCLLLPTLNEEQYGSLNEFFFCGEILPHRAAKALVSR
                     FPSATIYNTYGPTEATVAVTSIQITQEILDQYPTLPVGVERLGARLSTTDDGELVIEG
                     QSVSLGYLKNDQKTAEVFNFDDGIRTYHTGDKAKFENGQWFIQGRIDFQIKLNGYRME
                     LEEIETQLRQSEFVKEAIVVPVYKNDKVIHLIGAIVPTTEVTDNAEMTKNIKNDLKSR
                     LPEYMIPRKFEWMEQLPLTSNGKIDRKKIAEVING"
     gene            1902. .3116
                     /gene="dltB"
     CDS             1902. .3116
                     /gene="dltB"
```

```
                     /note="ORFID:SA0794"
                     /codon_start=1
                     /transl_table=11
                     /product="DltB membrane protein"
                     /protein_id="BAB42033.1"
                     /db_xref="GI:13700737"
                     /translation="MIPYGDFTFFLIALIALLPVIILGFLGKRSYIYNGVVTAFMIVL
                     IFSSDKHNLFDQKYLSVQLISFIIYVVWQVLLIMFYYHSKPKNNSFSKFVTVMVLSIL
                     PLALVKVLQSTWLGGHQIHFHESKLIEFVGFLGISYVTFKSVQLIMEIRDGSIKEIKV
                     WKLIQFISFFPTISSGPIDRYKRFVKDDKKVPTGNEYRELVLKAIHMIMLGFLYKYIV
                     AYFINTYAIMPLQLDLHGFVNLWLYMYAYSLYLFFDFAGYSLFAIAFSYLFGIKTPPN
                     FDKPFKAKNIKDFWNRWHMTLSFWFRDCIYMRSLFYMSRKKLLKSQFAMSNVAFLINF
                     FIMGIWHGIEVYYIVYGLYHAALFIGYGYYERWRKKHPPRWQNGFTTALSIVITFHFV
                     TFGFLIFSGKLI"
     gene            3134. .3370
                     /gene="dltC"
     CDS             3134. .3370
                     /gene="dltC"
                     /note="ORFID:SA0795"
                     /codon_start=1
                     /transl_table=11
                     /product="D-alanyl carrier protein"
                     /protein_id="BAB42034.1"
                     /db_xref="GI:13700738"
                     /translation="MEFREQVLNLLAEVAENDIVKENPDVEIFEEGIIDSFQTVGLLL
                     EIQNKLDIEVSIMDFDRDEWATPNKIVEALEELR"
     gene            3367. .4542
                     /gene="dltD"
     CDS             3367. .4542
                     /gene="dltD"
                     /note="(glycerophosphate chain)
                     ORFID:SA0796"
                     /codon_start=1
                     /transl_table=11
                     /product="poly D-alanine transfer protein"
                     /protein_id="BAB42035.1"
                     /db_xref="GI:13700739"
                     /translation="MKLKPFLPILISGAVFIVFLLLPASWFTGLVNEKTVEDNRTSLT
                     DQVLKGTLIQDKLYESNKYYPIYGSSELGKDDPFNPAIALNKHNANKKAFLLGAGGST
                     DLINAVELASQYDKLKGKKLTFIISPQWFTNHGLTNQNFDARMSQTQINQMFQQKNMS
                     TELKRRYAQRLLQFPHVHNKEYLKSYAKNPKETKDSYISGFKENQLIKIEAIKSLFAM
                     DKSPLEHVKPATKPDASWDEMKQKAVEIGKADTTSNKFGIRDQYWILIQESKRKVRRD
                     YEFNVNSPEFQDLELLVKTMRAAGADVQYVSIPSNGVWYDHIGIDKERRQAVYKKIHS
                     TVVDNGGKIYDMTDKDYEKYVISDAVHIGWKGWVYMDEQIAKHMKGEPQPEVDKPKN"
     gene            complement(4805. .5047)
                     /gene="SA0797"
     CDS             complement(4805. .5047)
                     /gene="SA0797"
                     /note="ORFID:SA0797
                     hypothetical protein, similar to nitrogen fixation protein
                     NifU"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB42036.1"
                     /db_xref="GI:13700740"
                     /translation="MPTEDTTMFDQVAEVIERLRPFLLRDGGDCSLIDVEDGIVKLQL
                     HGACGTCPSSTITLKAGIERALHEEVPGVIEVEQVF"
     gene            5163. .5471
                     /gene="SA0798"
     CDS             5163. .5471
                     /gene="SA0798"
                     /note="ORFID:SA0798"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42037.1"
                     /db_xref="GI:13700741"
                     /translation="MVVYGADVICASCVNAPTSKDIYDWLQPLLKRKYPNISFKYTYI
                     DITKDNDNLTDHDLQFIERIEQDELFYPLITMNDEYVADGYIQTKQITRFIDQKLVNE
                     "
     gene            complement(5530. .6594)
                     /gene="SA0799"
     CDS             complement(5530. .6594)
```

```
                     /gene="SA0799"
                     /note="ORFID:SA0799
                     hypothetical protein, similar to NADH dehydrogenase"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB42038.1"
                     /db_xref="GI:13700742"
                     /translation="MKNLVLLGGGYGNMRIMSRILTTSLPQDYTVTLVDRMPFHGLKP
                     EFYALAAGTKSDKDVRMKFPNHPQVNTVYGEINDIDLDAQIVSVGNSKIDYDELIIGL
                     GCEDKYHNVPGAEEYTHSIQTLSKARDTFHSISELPEGAKVGIVGAGLSGIELASELR
                     ESRSDLEIYLYDRGPRILRNFPEKLSKYVAKWFAKNNVTVVPNSNINKVEPGKIYNCD
                     EPKDIDLVVWTAGIQPVEVVRNLPIDINSNGRVIVNQYHQVPTYRNVYVVGDCADLPH
                     APSAQLAEVQGDQIADVLKKQWLNEPLPDKMPELKVQGIVGSLGDKQGFAYIMDRTVT
                     GRLASILKSGVLWLYKYHNG"
     gene            6911. .7147
                     /gene="SA0800"
     CDS             6911. .7147
                     /gene="SA0800"
                     /note="ORFID:SA0800"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42039.1"
                     /db_xref="GI:13700743"
                     /translation="MNPIVEFCLSNMAKGGDYVFNQLENDPDVDVLEYGCLTHCGICS
                     AGLYALVNGDIVEGDSPEELLQNIYAHIKETWIF"
     gene            7160. .7519
                     /gene="SA0801"
     CDS             7160. .7519
                     /gene="SA0801"
                     /note="ORFID:SA0801"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB42040.1"
                     /db_xref="GI:13700744"
                     /translation="MPTVILTEAAAYEVKDMLKANEMPDGYLKIKVNGGGCTGLTYGM
                     GAEEAPGENDEVLEYFGLKVLVDKEDAPVLNGTTIDFKQSLMGGGFQIDNPNAIASCG
                     CGSSFRTAKVAGNPENC"
     gene            7973. .9181
                     /gene="SA0802"
     CDS             7973. .9181
                     /gene="SA0802"
                     /note="ORFID:SA0802
                     hypothetical protein, similar to NADH dehydrogenase"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB42041.1"
                     /db_xref="GI:13700745"
                     /translation="MAQDRKKVLVLGAGYAGLQTVTKLQKAISTEEAEITLINKNEYH
                     YEATWLHEASAGTLNYEDVLYPVESVLKKDKVNFVQAEVTKIDRDAKKVETNQGIYDF
                     DILVVALGFVSETFGIEGMKDHAFQIENVITARELSRHIEDKFANYAASKEKDDNDLS
                     ILVGGAGFTGVEFLGELTDRIPELCSKYGVDQNKVKITCVEAAPKMLPMFSEELVNHA

  Query Match             26.7%;  Score 269.2;  DB 1;  Length 298050;
  Best Local Similarity   55.4%;  Pred. No. 2.9e-32;
  Matches  545;  Conservative    0;  Mismatches  433;  Indels    6;  Gaps    1;

Qy          17 GAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCAT 76
               |||||  | |  |    ||    || |    |||  |||| || ||   ||    ||
Db      113840 GAGGTAAATATGTGAATAGGAATATCGTTAAATTAGTTGTGTTTATGCTAATCTTAGTTG 113899

Qy          77 GTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCA 136
                 | |       |||   || |   |  ||  ||      ||  ||         ||| |
Db      113900 TAGCAGTAGCGGGTTGTGGTCAAAAAGATACTGAAGAGAAAACTGAAATGACGACAATAA 113959

Qy         137 AGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGT 196
               |  | |||  ||| |||   ||| |    ||| | |||||||||||||||||  | || |
Db      113960 AAGATGAATTAGGAACTGAAAAAATTAAGAAAAATCCTAAACGTGTTGTTGTATTAGAAT 114019

Qy         197 ATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
               ||   |||| ||||   ||||  || |||||| | ||||||||||| || || ||||||
Db      114020 ATAGTTTTGCTGATTATTTAGCAGCATTAGATATGAAACCTGTTGGTATTGCAGATGATG 114079

Qy         257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
                ||  |  ||   ||| |  ||  || |||||||||| |||||   |||    || || |
Db      114080 GCAGCACTAAAAATATAACAAAGTCAGTAAGAGATAAGATTGGGGCATATGAATCGGTTG 114139

Qy         317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
               || |  |   |||||| ||  | ||||  || |||||| | ||||| ||||| || ||||
Db      114140 GATCTAGACCGCAACCGAATATGGAAGTGATAAGTAAATTAAAACCGGATTTGATCATTG 114199

Qy         377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
               | |||  ||  ||||| ||    ||  |   ||| || |  ||||||||||| || ||
Db      114200 CAGATGTTAGCAGACATAAGAAAATCAAATCAGAATTGAGCAAAATTGCTCCGACAATCA 114259

Qy         437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
                  |    ||       ||||||||||||| ||||||||| || ||||||||| |  | |
Db      114260 TGTTAGTTAGCGGTACGGGAGATTATAATGCAAATATTGAAGCATTTAAAACAGTCGCTA 114319

Qy         497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
               ||||  |||| |||||  |||||||  | || ||  | ||| | || ||||| | | | |
Db      114320 AAGCAGTAGGCAAAGAGAAAGAAGGCGAGAAGCGTCTGGAAAAGCATGATAAAATATTAG 114379

Qy         557 AAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTG 616
                 ||   || |||  ||||      |  ||      ||   | |||  |    |
Db      114380 CGGAGATTAGAAAGAAAATTGAACAGAGTACGTTAAAATCTGCATTTGCATTCGGTATCT 114439

Qy         617 CTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAAC 676
               | | | ||||| || ||  | ||       |  | | | | || |||||| ||| | ||
Db      114440 CAAGAGCAGGTATGTTTATTAATAATGAAGATACATTTATGGGACAATTCTTAATTAAAA 114499

Qy         677 TAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAG 736
               | ||| || ||   | | | |   | ||    ||||    | |  || |          |
Db      114500 TGGGTATTCAACCTGAAGTCACAAAAGACAAAACTACGCATGTTGGTGAACGCAAGGGTG 114559

Qy         737 GACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCA 796
               | |||||  || |  | || | ||||    |  |  |  | |||||  |   |||| |
Db      114560 GTCCTTATATATATTTAAATAATGAAGAACTTGCCAATATCAATCCAAAAGTTATGATTT 114619

Qy         797 TAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTG 856
               ||   ||  ||  |  ||     ||  |         ||   || |     |   | | |
Db      114620 TAGCCACTGACGGAAAAACGGACAAAAA------TAGAACGAAATTCATTGATCCTGCAG 114673

Qy         857 TATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTAT 916
               | |||||   |||||| ||||||||| || |    ||| ||    | |||||  |    |
Db      114674 TTTGGAAATCATTAAAAGCTGTGAAAGATAACAAAGTTTATGACGTTGACCGAAATAAGT 114733

Qy         917 GGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTAT 976
               ||   | |||| | ||  | ||  |    |||   |||||| ||||  | |   || |
Db      114734 GGTTGAAATCAAGGGGGATTATCGCAAGTGAAAGTATGGCAGAAGATTTAGAAAAAATTG 114793

Qy         977 CTAAGAAAGATAGTAAAAAAGATA 1000
               |  | ||||  |   ||||| | |
Db      114794 CAGAAAAAGCAAAATAAAAATACA 114817


RESULT 10
AP003361
LOCUS       AP003361               347785 bp    DNA     linear   BCT 17-APR-2004
DEFINITION  Staphylococcus aureus subsp. aureus Mu50 DNA, complete genome,
            section 4/9.
ACCESSION   AP003361 BA000017
VERSION     AP003361.3  GI:46395333
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus Mu50
  ORGANISM  Staphylococcus aureus subsp. aureus Mu50
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
            Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
            Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
            Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
            Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
            Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
            Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
  TITLE     Whole genome sequencing of meticillin-resistant Staphylococcus
            aureus
  JOURNAL   Lancet 357 (9264), 1225-1240 (2001)
   MEDLINE  21311952
   PUBMED   11418146
REFERENCE   2
  AUTHORS   Ohta,T., Hirakawa,H., Morikawa,K., Maruyama,A., Inose,Y.,
            Yamashita,A., Oshima,K., Kuroda,M., Hattori,M., Hiramatsu,K.,
            Kuhara,S. and Hayashi,H.
  TITLE     Nucleotide Substitutions in Staphylococcus aureus Strains, Mu50,
            Mu3, and N315
  JOURNAL   DNA RESEARCH 11, 51-56 (2004)
REFERENCE   3  (bases 1 to 347785)
  AUTHORS   Hattori,M., Yamashita,A., Oshima,K. and Shiba,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2001) Masahira Hattori, Kitasato Institute for
            Life Sciences, Kitasato University; Kitasato 1-15-1, Sagamihara,
            Kanagawa 228-8555, Japan
            (E-mail:hattori@genome.ls.kitasato-u.ac.jp, Tel:81-42-778-8194,
            Fax:81-42-778-8193)
COMMENT     On Apr 15, 2004 this sequence version replaced gi:14246761.
            This work was done in collaboration with Toshiko Ohta, Mutsumi
            Kanamori, Hideo Hayashi (University of Tsukuba), Hideki Hirakawa,
            Satoru Kuhara (Kyushu University), Ikuo Uchiyama (Okazaki National
            Research Institutes), Susumu Goto, Minoru Kanehisa (Kyoto
            University), Naotake Ogasawara (Nara Institute of Science and
            Technology) and supported by the Research for the Future Program of
            the Japan Society for the Promotion of Science.
FEATURES             Location/Qualifiers
     source          1. .347785
                     /organism="Staphylococcus aureus subsp. aureus Mu50"
                     /mol_type="genomic DNA"
                     /strain="Mu50"
                     /sub_species="aureus Mu50"
                     /db_xref="taxon:158878"
     gene            11. .997
                     /locus_tag="SAV0992"
     CDS             11. .997
                     /locus_tag="SAV0992"
                     /codon_start=1
                     /transl_table=11
                     /product="oligopeptide ABC transporter ATP-binding protein
                     homologue"
                     /protein_id="BAB57154.1"
                     /db_xref="GI:14246762"
                     /translation="MNNVLLEVKDLETSLKINNEWLATVENISFELSKGEVLGIVGES
                     GCGKSILSKSIIKLLPEKISKLSSGEVIFDGKRIDTLNEKQLLDIRGNDIAMIFQEPM
                     TALNPVFTIKNQLVESIKSHKKISKKEANNLAKDLLKKVGIARQDEILNSYPHQLSGG
                     MRQRVMIAMAISCSPKLLIADEPTTALDVTIQAQILDLLKELQKETQMAIMMITHDLS
                     VVAEFCDKVLVMYAGQIVEFGGIKEILHNPKHPYTQKLLSTIPKLKEGQKRLETIEGI
                     VPSIQAFHVNKCRFANRCNKKLDICNNQSPKMHVCEDVIVRCHLYKNEYKEI"
     gene            1000. .1980
                     /gene="appF"
                     /locus_tag="SAV0993"
     CDS             1000. .1980
                     /gene="appF"
                     /locus_tag="SAV0993"
                     /codon_start=1
                     /transl_table=11
                     /product="oligopeptide transport system ATP-binding
                     protein AppF homologue"
                     /protein_id="BAB57155.1"
                     /db_xref="GI:14246763"
                     /translation="MENILEVNQIKKYYKIKTGLLQKTQYVKAVDDVSFSIKKGQTFG
                     LVGESGCGKSTLGKVIIRLEDATSGSIIVNGEDITRLQGKKLRKSRQQYQMIFQDPYA
                     SLNPMQMVGDIISEPILNYKKLPKEEIKKEVLYLLKCVGLSEDAYYKYAHEFSGGQRQ
                     RVGIARALALRPSLIVADEPVSALDVSVQSQVLNLLKDLQEQFKLSYLFIAHDLSVVK
                     HISDVIGVMYLGHIVEIASDKEIYENPKHPYTKALISSIPQIDKHNNNRIILKGELPS
                     PSNPPQGCPFHTRCPIAKDKCKENIPQLKDIGNEHQVACFYVDKVGDLNA"
     gene            1973. .2935
                     /gene="oppB"
                     /locus_tag="SAV0994"
     CDS             1973. .2935
                     /gene="oppB"
                     /locus_tag="SAV0994"
                     /codon_start=1
                     /transl_table=11
                     /product="probable oligopeptide transport system permease
                     protein"
                     /protein_id="BAB57156.1"
                     /db_xref="GI:14246764"
                     /translation="MLKLILKRLGLMIPLLILISIVVFSLAIIQPGDPFSDLQNGKIK
                     QEAINAQREKLGLNDSIPHQYIRWVNHVIHGDLGESIKYKRPVIDVIEERIPNTILLG
                     AMSLIITYIISFALGITSGRYSYSLTDYTVQIFNYLMLAIPSFIAGVFAIFIFSFELQ
                     WFPFQGSVDINLKEGTFEYYMSKIYHTFLPALTLGLLSTAGYIQYLRNDIIENSKKDY
                     VLTARSKGLSMNKIYNKHILRNSLIPIITFLGADIVSILGGAVITETIFSYNGIGKLF
                     LESVIGQDYPLMMALTLFFSFLGLLGNLISDITYGFIDPRIRSN"
     gene            2947. .3828
                     /locus_tag="SAV0995"
     CDS             2947. .3828
                     /locus_tag="SAV0995"
                     /codon_start=1
                     /transl_table=11
                     /product="similar to oligopeptide ABC transporter
                     permease"
                     /protein_id="BAB57157.1"
                     /db_xref="GI:14246765"
                     /translation="MQNKSKSPFKIALSRFIHNKIAMLSIIFLLIITIVSIIAPLIAP
                     FPVNQQDLLNIKGEMTAQNILGTDSGGRDNFSRLLYAGRISLSIGITSTIGMLLIGIT
                     VGVISGYFGGIVDTLLMRITEFVMLFPFLIFAIVLNAALGDKIKNPYGSAIILVLVII
                     VLSWGGIARLVRGKVLQEKENEYFLAAKSIGTPTYKIILKHLLPNILSVVIVQATLLF
                     AGMIVVESGLSFLGFGISKAIPSWGNMLSDAQEGDVISGKPWIWMPPAIMITLTILSI
                     NFVGEGLKDAFNPRGRR"
     gene            complement(3870. .4859)
                     /gene="trpS"
                     /locus_tag="SAV0996"
     CDS             complement(3870. .4859)
                     /gene="trpS"
                     /locus_tag="SAV0996"
                     /codon_start=1
                     /transl_table=11
                     /product="tryptophanyl-tRNA synthetase"
                     /protein_id="BAB57158.1"
                     /db_xref="GI:14246766"
                     /translation="METLFSGIQPSGIPTIGNYIGALKQFVDVQNDYDCYFCIVDQHA
                     ITMPQDRLKLRKQTRQLAAIYLASGIDPDKATLFIQSEVPAHVQAGWMLTTIASVGEL
                     ERMTQYKDKAQKAVEGIPAGLLTYPPLMAADIVLYNTNIVPVGDDQKQHIELTRNLVD
                     RFNSRYNDVLVKPEIRMPKVGGRVMSLQDPTRKMSKSDDNAKNFISLLDEPNVAAKKI
                     KSAVTDSDGIIKFDRDNKPGITNLISIYAGLTDMPIKDIEAKYEGEGYGKFKGDLAEI
                     VKAFLVEFQEKYESFYNSDKLDDILDQGRDKAHKVSFKTVKKMEKAMGLGRKR"
     gene            5153. .5548
                     /locus_tag="SAV0997"
     CDS             5153. .5548
                     /locus_tag="SAV0997"
                     /codon_start=1
                     /transl_table=11
                     /product="arsenate reductase family protein"
                     /protein_id="BAB57159.1"
                     /db_xref="GI:14246767"
                     /translation="MVTLFTSPSCTSCRKAKAWLQEHDIPYTERNIFSEHLTIDEIKQ
                     ILKMTEDGTDEIISTRSKTYQKLNVDIDSLPLQDLYSIIQDNPGLLRRPIILDNKRLQ
                     VGYNEDEIRRFLPRKVRTFQLQEAQRMVD"
     gene            5919. .6638
                     /locus_tag="SAV0998"
     CDS             5919. .6638
                     /locus_tag="SAV0998"
                     /codon_start=1
                     /transl_table=11
                     /product="similar to negative regulator of genetic
                     competence MecA"
                     /protein_id="BAB57160.1"
                     /db_xref="GI:14246768"
                     /translation="MRIERVDDTTVKLFITYSDIEARGFSREDLWTNRKRGEEFFWSM
                     MDEINEEEDFVVEGPLWIQVHAFEKGVEVTISKSKNEDMMNMSDDDATDQFDEQVQEL
                     LAQTLEGEDQLEELFEQRTKEKEAQGSKRQKSSARKNTRTIIVKFNDLEDVINYAYHS
                     NPITTEFEDLLYMVDGTYYYAVHFDSHVDQEVINDSYSQLLEFAYPTDRTEVYLNDYA
                     KIIMSHNVTAQVRRYFPETTE"
     gene            6759. .7745
                     /locus_tag="SAV0999"
     CDS             6759. .7745
                     /locus_tag="SAV0999"
                     /codon_start=1
                     /transl_table=11
                     /product="similar to transcription factor"
                     /protein_id="BAB57161.1"
                     /db_xref="GI:14246769"
                     /translation="MLVALNEEKERVLATTALRKTQYFCPVCGKQVILKRGLKVISHF
                     AHKHLAEQKCFNNETIKHYKSKLILAQMIQQQGCKVEIEPFLKEIKQIPDILINNKYV
                     IELQYSPIPYKQILQRTEGLKKMGYKVSWLLNDVDYCHNKVKFNHFHSLFINPITRKL
                     HTFNLEKKQIMMFQQIQYLGGHKYVAEKRNAKIIELFNEAPCDYHAVYKLSKFAINQY
                     IKYCRWQNSVLEPTLSAMYQLQLTDQEVVYNYGYIFPEQIYIENHPIEWQLQVDLWLK
                     NGKSKLVNDNLNYFKLKKFIVALESKTAIIEKLINNYLNICSDRGNDVQILF"
     gene            7793. .9601
                     /locus_tag="SAV1000"
     CDS             7793. .9601
                     /locus_tag="SAV1000"
                     /codon_start=1
                     /transl_table=11
                     /product="thimet oligopeptidase homologue"
                     /protein_id="BAB57162.1"
                     /db_xref="GI:14246770"
                     /translation="MSQQLSREEQERKYPEYTWDLTTIFKDDEAFEAAFKEVENELGK
                     EEQFKGHIGDSAETLYNALELEDTLGTKLEKVYVYAHLKQDQDTTNDKYTGMESRAHQ
                     LIIKFSSAWSFLVPEILQIDEDKIQSFVNSYDKLQKFAFDLKLINEKRPHILDAETEK
                     LLTEAQDALSTPSNVYGMFSNADLVFEDAIDKDGNAHPLTQGTFIKYLESDDRKLRES
                     AFRNVYKAYGAHNNTLGATLAGEVKKNVFNARTHNYKTAREKALSNNHIPENVYDNLV
                     KTVHKYLPLLHRYTELRKELLGLDDLKMYDLYTPLIKDIKFEMPYEEAKEWMLKALEP
                     MGEEYLNVVKEGLNNRWVDVYENKGKRSGGYSSGAHLTNPFILLNWSNTISDLYTLVH
                     EFGHSAHSYFSRKFQPSNSSDYTIFVAEVASTCNEALLSDYMDKHLDDEKRLLLLNQE
                     LERFRATLFRQTMFAEFEHKIHAIEEAGEPLTPTRMNEEYAKLNKLYFGDSVETDEDI
                     SKEWSRIPHFYMNYYVYQYATGYSAAQSLSHQILTEGKPAVDRYINEFLKKGSSNYPI
                     EILKNAGVDMTTPEPIEQACEVFEQKLNAFEKLMKA"
     gene            complement(10061. .10846)

  Query Match             26.7%;  Score 269.2;  DB 1;  Length 347785;
  Best Local Similarity   55.4%;  Pred. No. 2.8e-32;
  Matches  545;  Conservative    0;  Mismatches  433;  Indels    6;  Gaps    1;

Qy          17 GAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCAT 76
               |||||  | |  |    ||    || |    |||  |||| || ||   ||    ||
Db       46202 GAGGTAAATATGTGAATAGGAATATCGTTAAATTAGTTGTGTTTATGCTAATCTTAGTTG 46261

Qy          77 GTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCA 136
                 | |       |||   || |   |  ||  ||      ||  ||         ||| |
Db       46262 TAGCAGTAGCGGGTTGTGGTCAAAAAGATACTGAAGAGAAAACTGAAATGACGACAATAA 46321

Qy         137 AGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGT 196
               |  | |||  ||| |||   ||| |    ||| | |||||||||||||||||  | ||  |
Db       46322 AAGATGAATTAGGAACTGAAAAAATTAAGAAAAATCCTAAACGTGTTGTTGTATTAGAAT 46381

Qy         197 ATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
               ||   |||| ||||   ||||  || |||||| | ||||||||||| || || ||||||
Db       46382 ATAGTTTTGCTGATTATTTAGCAGCATTAGATATGAAACCTGTTGGTATTGCAGATGATG 46441

Qy         257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
                ||  |  ||   ||| |  ||  || |||||||||| |||||   |||    ||  || |
Db       46442 GCAGCACTAAAAATATAACAAAGTCAGTAAGAGATAAGATTGGGGCATATGAATCGGTTG 46501

Qy         317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
               || |  |   |||||| ||  | ||||  || |||||| | ||||| ||||| || ||||
Db       46502 GATCTAGACCGCAACCGAATATGGAAGTGATAAGTAAATTAAAACCGGATTTGATCATTG 46561

Qy         377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
               | |||  ||  ||||| ||    ||  |   ||| || |  ||||||||||| || ||
Db       46562 CAGATGTTAGCAGACATAAGAAAATCAAATCAGAATTGAGCAAAATTGCTCCGACAATCA 46621

Qy         437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
                  |    ||       ||||||||||||| ||||||||| || ||||||||| |  | |
Db       46622 TGTTAGTTAGCGGTACGGGAGATTATAATGCAAATATTGAAGCATTTAAAACAGTCGCTA 46681

Qy         497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
               ||||  |||| |||||  |||||||  | || ||  | ||| | || ||||| | | | |
Db       46682 AAGCAGTAGGCAAAGAGAAAGAAGGCGAGAAGCGTCTGGAAAAGCATGATAAAATATTAG 46741

Qy         557 AAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTG 616
                 ||   || |||  ||||      |  ||      ||   | |||  |    |
Db       46742 CGGAGATTAGAAAGAAAATTGAACAGAGTACGTTAAAATCTGCATTTGCATTCGGTATCT 46801

Qy         617 CTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAAC 676
               | | | ||||| || ||  | ||       |  | | | | || |||||| ||| | ||
Db       46802 CAAGAGCAGGTATGTTTATTAATAATGAAGATACATTTATGGGACAATTCTTAATTAAAA 46861

Qy         677 TAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAG 736
               | ||| || ||   | | | |   | ||    ||||    | |  || |          |
Db       46862 TGGGTATTCAACCTGAAGTCACAAAAGACAAAACTACGCATGTTGGTGAACGCAAGGGTG 46921

Qy         737 GACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCA 796
               | |||||  || |  | || | ||||    |  |  |  | |||||  |   |||| |
Db       46922 GTCCTTATATATATTTAAATAATGAAGAACTTGCCAATATCAATCCAAAAGTTATGATTT 46981

Qy         797 TAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTG 856
               ||   ||  ||  |  ||     ||  |         ||    || |     |   | | |
Db       46982 TAGCCACTGACGGAAAAACGGACAAAAA------TAGAACGAAATTCATTGATCCTGCAG 47035

Qy         857 TATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTAT 916
               | |||||   |||||| ||||||||| || |    ||| ||    | |||||  |    |
Db       47036 TTTGGAAATCATTAAAAGCTGTGAAAGATAACAAAGTTTATGACGTTGACCGAAATAAGT 47095

Qy         917 GGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTAT 976
               ||   | |||| | ||  | ||  |    |||   |||||| ||||  | |   || |
Db       47096 GGTTGAAATCAAGGGGGATTATCGCAAGTGAAAGTATGGCAGAAGATTTAGAAAAAATTG 47155

Qy         977 CTAAGAAAGATAGTAAAAAAGATA 1000
               |  | ||||  |   ||||| | |
Db       47156 CAGAAAAAGCAAAATAAAAATACA 47179


RESULT 11
AX621226
LOCUS       AX621226                 975 bp    DNA     linear   PAT 20-FEB-2003
DEFINITION  Sequence 4189 from Patent WO02094868.
ACCESSION   AX621226
VERSION     AX621226.1  GI:28449932
KEYWORDS    .
SOURCE      Staphylococcus aureus
  ORGANISM  Staphylococcus aureus
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Masignani,V.C., Mora,M.C. and Scarselli,M.C.
  TITLE     Staphylococcus aureus proteins and nucleic acids
  JOURNAL   Patent: WO 02094868-A 4189 28-NOV-2002;
            Chiron Spa (IT)
FEATURES             Location/Qualifiers
     source          1. .975
                     /organism="Staphylococcus aureus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1280"
ORIGIN

  Query Match             26.2%;  Score 263.8;  DB 6;  Length 975;
  Best Local Similarity   55.3%;  Pred. No. 8.3e-31;
  Matches  536;  Conservative    0;  Mismatches  427;  Indels    6;  Gaps    1;

Qy          17 GAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCAT 76
               |||||  |    |    ||    || |    | |  |||| || ||   || |  ||
Db           8 GAGGTAAATGTGTGAATAGGAATATCGTTAAACTAGTTGTGTTCATGCTAATTTTAGTTG 67

Qy          77 GTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCA 136
                 | |       |||   || |   |  ||  ||      ||  ||         ||| |
Db          68 TAGCAGTAGCGGGTTGTGGTCAAAAAGATACTGAAGAGAAAACTGAAATGACGACAATAA 127

Qy         137 AGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGT 196
```

```
              |  | |||  ||| |||   ||| |    ||| | |||||||||||||||||  | || |
Db        128 AAGATGAATTAGGAACTGAAAAAATTAAGAAAAATCCTAAACGTGTTGTTGTATTAGAAT 187

Qy        197 ATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
              ||   |||| ||||   |||| |  || |||||| | ||||||||||||| || || ||||||
Db        188 ATAGTTTTGCTGATTATTTAGCAGCATTAGATATGAAACCTGTTGGTATTGCAGATGATG 247

Qy        257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
               ||  |  ||   ||| |  ||  || |||||||||| |||||   |||    || || |
Db        248 GCAGCACTAAAAATATAACAAAGTCAGTAAGAGATAAGATTGGGGCATATGAATCGGTTG 307

Qy        317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
              || |  |   |||||| ||  | ||||  || |||||| | ||||| |||||| || ||||
Db        308 GATCTAGACCGCAACCGAATATGGAAGTGATAAGTAAATTAAAACCGGATTTGATCATTG 367

Qy        377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
              | |||  ||  ||||| ||    ||  |   ||| || |  ||||||||||||| || ||
Db        368 CAGATGTTAGCAGACATAAGAAAATCAAATCAGAATTGAGCAAAATTGCTCCGACAATCA 427

Qy        437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
                 |    ||       ||||||||||||| ||||||||| || |||||||||| |  | |
Db        428 TGTTAGTTAGCGGTACGGGAGATTATAATGCAAATATTGAAGCATTTAAAACAGTCGCTA 487

Qy        497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
              ||||  |||| |||||  |||||||  | || | ||  | ||| | | || ||||| | | | |
Db        488 AAGCAGTAGGCAAAGAGAAAGAAGGCGAGAAGCGTCTGGAAAAGCATGATAAAATATTAG 547

Qy        557 AAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTG 616
                ||   || |||  ||||      |  ||      ||   | |||  |    |
Db        548 CGGAGATTAGAAAGAAAATTGAACAGAGTACGTTAAAATCTGCATTTGCATTCGGTATCT 607

Qy        617 CTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAAC 676
              | | | ||||| || ||  | ||       |  | | | | || |||||| ||| | ||
Db        608 CAAGAGCAGGTATGTTTATTAATAATGAAGATACATTTATGGGACAATTCTTAATTAAAA 667

Qy        677 TAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAG 736
              | ||| || ||   | | | | |   | ||    ||||     | |  || |          |
Db        668 TGGGTATTCAACCTGAAGTCACAAAAGACAAAACTACGCATGTTGGTGAACGCAAGGGTG 727

Qy        737 GACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCA 796
              | |||||  || |  | || | ||||    |  |  |  | |||||  |   |||| |
Db        728 GTCCTTATATATATTTAAATAATGAAGAACTTGCCAATATCAATCCAAAAGTTATGATTT 787

Qy        797 TAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTG 856
              ||   ||  ||  |  ||     ||  |         ||   || |     |   | | |
Db        788 TAGCCACTGACGGAAAAACGGACAAAAA------TAGAACGAAATTCATTGATCCTGCAG 841

Qy        857 TATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTAT 916
              | |||||   |||||| ||||||||| |  |    ||| ||    | |||||  |    |
Db        842 TTTGGAAATCATTAAAAGCTGTGAAAGATAACAAAGTTTATGACGTTGACCGAAATAAGT 901

Qy        917 GGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTAT 976
              ||   | |||| | ||  | ||  |    |||   ||||||| ||||  | |   || |
Db        902 GGTTGAAATCAAGGGGGATTATCGCAAGTGAAAGTATGGCAGAAGATTTAGAAAAAATTG 961

Qy        977 CTAAGAAAG 985
              |  | ||||
Db        962 CAGAAAAAG 970

RESULT 12
AX583924
LOCUS       AX583924                 957 bp    DNA     linear   PAT 10-JAN-2003
DEFINITION  Sequence 301 from Patent WO02059148.
ACCESSION   AX583924
VERSION     AX583924.1  GI:27655594
KEYWORDS    .
SOURCE      Staphylococcus aureus
  ORGANISM  Staphylococcus aureus
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Meinke,A., Nagy,E., von Ahsen,U., Klade,C., Henics,T., Zauner,W.,
            Minh,D.B., Vytvytska,O., Etz,H., Dryla,A., Weichhart,T., Hafner,M.,
            Tempelmaier,B., Fraser,C.M. and Gill,S.
  TITLE     A method for identification, isolation and production of antigens
            to a specific pathogen
  JOURNAL   Patent: WO 02059148-A 301 01-AUG-2002;
            Cistem Biotechnologies GmbH (AT)
FEATURES             Location/Qualifiers
     source          1. .957
                     /organism="Staphylococcus aureus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1280"
ORIGIN

  Query Match             26.1%;  Score 263.2;  DB 6;  Length 957;
  Best Local Similarity   55.9%;  Pred. No. 1e-30;
  Matches  524;  Conservative    0;  Mismatches  408;  Indels    6;  Gaps    1;

Qy         48 ATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAA 107
              | |  |||| || ||   || |  ||     | |       |||   || |   |  ||
Db         21 ACTAGTTGTGTTCATGCTAATTTTAGTTGTAGCAGTAGCGGGTTGTGGTCAAAAAGATAC 80

Qy        108 AGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAA 167
               ||      ||  ||         ||| ||  | |||  |||| |||   ||| |    ||
Db         81 TGAAGAGAAAACTGAAATGACGACAATAAAAGATGAATTAGGAACTGAAAAAATTAAGAA 140

Qy        168 ACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGA 227
              | | ||||||||||||||||||  | || |||   |||| ||||   ||||  || |||||
Db        141 AAATCCTAAACGTGTTGTTGTATTAGAATATAGTTTTGCTGATTATTTAGCAGCATTAGA 200

Qy        228 TGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAG 287
              | | |||||||||||| || || ||||||  ||  |  ||   ||| |  ||  || ||||
Db        201 TATGAAACCTGTTGGTATTGCAGATGATGGCAGCACTAAAAATATAACAAAGTCAGTAAG 260

Qy        288 AGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAAT 347
              |||||| |||||   |||    || || ||| |  |   |||||| ||  | ||||  ||
Db        261 AGATAAGATTGGGGCATATGAATCGGTTGGATCTAGACCGCAACCGAATATGGAAGTGAT 320

Qy        348 CAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAA 407
               |||||| | ||||| |||||| || ||||| |||  ||  ||||| ||    ||  |
Db        321 AAGTAAATTAAAACCGGATTTGATCATTGCAGATGTTAGCAGACATAAGAAAATCAAATC 380

Qy        408 AGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGA 467
              ||| || |  ||||||||||| || ||     |    ||       |||||||||||||
Db        381 AGAATTGAGCAAAATTGCTCCGACAATCATGTTAGTTAGCGGTACGGGAGATTATAATGC 440

Qy        468 AAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAA 527
              ||||||||| || ||||||||||| |  | ||||||  |||| |||||  |||||||  | ||
Db        441 AAATATTGAAGCATTTAAAACAGTCGCTAAAGCAGTAGGCAAAGAGAAAGAAGGCGAGAA 500

Qy        528 ACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAA 587
               ||  | ||| | || |||||| | | | |  ||   || |||  ||||      |  ||
Db        501 GCGTCTGGAAAAGCATGATAAAATATTAGCGGAGATTAGAAAGAAAATTGAACAGAGTAC 560

Qy        588 AAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAA 647
                   ||   | |||  |    |      | | | ||||| || ||  | ||       |
Db        561 GTTAAAATCTGCATTTGCATTCGGTATCTCAAGAGCAGGTATGTTTATTAATAATGAAGA 620

Qy        648 CTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGT 707
                | | | | || |||||| ||| | || | ||| || ||   | | | |   | ||
Db        621 TACATTTATGGGACAATTCTTAATTAAAATGGGTATTCAACCTGAAGTCACAAAAGACAA 680

Qy        708 TACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTT 767
               ||||    | |  || |          || |||||  || |  | || | ||||    |
Db        681 AACTACGCATGTTGGTGAACGCAAGGGTGGTCCTTATATATATTTAAATAATGAAGAACT 740

Qy        768 ATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACC 827
                |  |  | |||||  |   |||| |  ||   ||  ||  |  ||     ||  |
Db        741 TGCCAATATCAATCCAAAAGTTATGATTTTAGCCACTGACGGAAAAACGGACAAAAA--- 797

Qy        828 TTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCA 887
                    ||   || |     |   | | || |||||   ||||| |||||||||| || |
Db        798 ---TAGAACGAAATTCATTGATCCTGCAGTTTGGAAATCATTAAAAGCTGTGAAAGATAA 854

Qy        888 ACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGA 947
```

```
                   |||  ||    | |||||  |    |||   | |||| | ||  | ||  |    ||
Db         855 CAAAGTTTATGACGTTGACCGAAATAAGTGGTTGAAATCAAGGGGGATTATCGCAAGTGA 914

Qy         948 AGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAG 985
               |   |||||| ||||  | |   || |  |  | ||||
Db         915 AAGTATGGCAGAAGATTTAGAAAAAATTGCAGAAAAAG 952


RESULT 13
BX571857_10
WPCOMMENT
  Sequence split into 28 fragments  LOCUS BX571857 Accession BX571857
     Fragment Name        Begin        End
     BX571857_00              1     110000
     BX571857_01         100001     210000
     BX571857_02         200001     310000
     BX571857_03         300001     410000
     BX571857_04         400001     510000
     BX571857_05         500001     610000
     BX571857_06         600001     710000
     BX571857_07         700001     810000
     BX571857_08         800001     910000
     BX571857_09         900001    1010000
     BX571857_10        1000001    1110000
     BX571857_11        1100001    1210000
     BX571857_12        1200001    1310000
     BX571857_13        1300001    1410000
     BX571857_14        1400001    1510000
     BX571857_15        1500001    1610000
     BX571857_16        1600001    1710000
     BX571857_17        1700001    1810000
     BX571857_18        1800001    1910000
     BX571857_19        1900001    2010000
     BX571857_20        2000001    2110000
     BX571857_21        2100001    2210000
     BX571857_22        2200001    2310000
     BX571857_23        2300001    2410000
     BX571857_24        2400001    2510000
     BX571857_25        2500001    2610000
     BX571857_26        2600001    2710000
     BX571857_27        2700001    2799802
Continuation (11 of 28) of BX571857 from base 1000001 (BX571857 Staphylococcus aureus st

  Query Match             26.1%;  Score 262.8;  DB 1;  Length 110000;
  Best Local Similarity   55.0%;  Pred. No. 3.6e-31;
  Matches  541;  Conservative    0;  Mismatches  437;  Indels    6;  Gaps    1;

Qy          17 GAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCAT 76
               ||||| ||    |    ||    || |    | |  |||| || ||   ||    ||
Db       42307 GAGGTATATGTGTGAATAGGAATATCGTTAAACTAGTTGTGTTTATGCTAATCTTAGTTG 42366

Qy          77 GTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCA 136
                 | |       |||   || |   |  ||  ||      ||  ||         ||| |
Db       42367 TAGCAGTAGCGGGTTGTGGTCAAAAAGATACTGAAGAGAAAACTGAAATGACGACAATAA 42426

Qy         137 AGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGT 196
               |  | |||  ||| |||   ||| |    ||| | |||||||||||||||||  | || |
Db       42427 AAGATGAATTAGGAACTGAAAAAATTAAGAAAAATCCTAAACGTGTTGTTGTATTAGAAT 42486

Qy         197 ATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
               ||   |||| ||||   ||||  || |||||| | ||||||||||| || || ||||||
Db       42487 ATAGTTTTGCTGATTATTTAGCAGCATTAGATATGAAACCTGTTGGTATTGCAGATGATG 42546

Qy         257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
                ||  |  ||   ||| |  ||  || |||||||||| |||||   |||    || || |
Db       42547 GCAGCAGTAAAAATATAACAAAGTCAGTAAGAGATAAGATTGGGGCATATGAATCGGTTG 42606

Qy         317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
               || |  |   |||||| ||  | ||||  || |||||| | ||||| ||||| || ||||
Db       42607 GATCTAGACCGCAACCGAATATGGAAGTGATAAGTAAATTAAAACCGGATTTGATCATTG 42666

Qy         377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
               | |||  ||  ||||| ||    ||  |   ||| || |  |||||||| || || |||
Db       42667 CAGATGTTAGCAGACATAAGAAAATCAAATCAGAATTGAGCAAAATTGCGCCGACAATTA 42726

Qy         437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
                  |    |||      |||||||| |||| ||||||||| || ||||||||| |  | |
Db       42727 TGTTAGTCAGTGGTACGGGAGATTACAATGCAAATATTGAAGCATTTAAAACAGTCGCTA 42786

Qy         497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
               ||||  | || |||||  |||||||  | || ||  | ||| | || ||||| | | | |
Db       42787 AAGCAGTTGGCAAAGAGAAAGAAGGCGAGAAGCGTCTGGAAAAGCATGATAAAATATTAG 42846

Qy         557 AAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTG 616
                 ||   || |||  ||||      |  ||      ||   | |||  |    |
Db       42847 CGGAGATTAGAAAGAAAATTGAACAGAGTACGTTAAAATCTGCATTTGCATTCGGTATCT 42906

Qy         617 CTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAAC 676
               | | | ||||| || ||  | ||       |  | | | | || |||||| ||  | ||
Db       42907 CAAGAGCAGGTATGTTTATTAATAATGAAGATACATTTATGGGACAATTCTTACTTAAAA 42966

Qy         677 TAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAG 736
               | ||| || ||   | | | |   | ||    ||||    | |  || |          |
Db       42967 TGGGTATTCAACCTGAAGTCACAAAAGACAAAACTACGCATGTTGGTGAACGCAAGGGTG 43026

Qy         737 GACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCA 796
               | |||||  || |  | || | ||||    |  |  |  | |||||  |   |||| |
Db       43027 GTCCTTATATATATTTAAATAATGAAGAACTTGCCAATATCAATCCAAAAGTTATGATTT 43086

Qy         797 TAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTG 856
               ||   ||  ||     ||     ||  |         ||   || |     |   | | |
Db       43087 TAGCCACTGACGGGAAAACGGACAAAAA------TAGAACGAAATTCATTGATCCTGCAG 43140

Qy         857 TATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTAT 916
               | |||||   |||||| ||||||||| || |    ||| ||    | |||||  |    |
Db       43141 TTTGGAAATCATTAAAAGCTGTGAAAGATAACAAAGTTTATGACGTTGACCGAAATAAGT 43200

Qy         917 GGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTAT 976
               ||   | |||| | ||  | ||  |    |||   |||||| ||||  | |   || |
Db       43201 GGTTGAAATCAAGGGGGATTATCGCAAGTGAAAGTATGGCAGAAGATTTAGAAAAAATTG 43260

Qy         977 CTAAGAAAGATAGTAAAAAAGATA 1000
               |  | ||||  |   ||||| | |
Db       43261 CAGAAAAAGCAAAATAAAAATACA 43284


RESULT 14
AP004825
LOCUS       AP004825               290150 bp    DNA     linear   BCT 20-DEC-2002
DEFINITION  Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,
            strain:MW2, section 4/10.
ACCESSION   AP004825 BA000033
VERSION     AP004825.1  GI:21203989
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus MW2
  ORGANISM  Staphylococcus aureus subsp. aureus MW2
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Baba,T., Takeuchi,F., Kuroda,M., Yuzawa,H., Aoki,K., Oguchi,A.,
            Nagai,Y., Iwama,N., Asano,K., Naimi,T., Kuroda,H., Cui,L.,
            Yamamoto,K. and Hiramatsu,K.
  TITLE     Genome and virulence determinants of high virulence
            community-acquired MRSA
  JOURNAL   Lancet 359 (9320), 1819-1827 (2002)
  MEDLINE   22040717
   PUBMED   12044378
REFERENCE   2  (bases 1 to 290150)
  AUTHORS   Director-General, Biotechnology Center, Aoki,K., Oguchi,A.,
            Nagai,Y., Asano,K., Iwama,N., Baba,T., Kuroda,M., Hiramatsu,K. and
            Kikuchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAR-2002) Director-General, Biotechnology Center,
            National Institute of Technology and Evaluation, Biotechnology
            Center; 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
            (E-mail:bio@nite.go.jp, URL:http://www.bio.nite.go.jp/,
            Tel:81-3-3481-1933, Fax:81-3-3481-8424)
```

```
FEATURES             Location/Qualifiers
     source          1. .290150
                     /organism="Staphylococcus aureus subsp. aureus MW2"
                     /mol_type="genomic DNA"
                     /strain="MW2"
                     /db_xref="taxon:196620"
     gene            285. .1493
                     /gene="MW0823"
     CDS             285. .1493
                     /gene="MW0823"
                     /note="ORFID:MW0823
                     hypothetical protein, similar to NADH dehydrogenase"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAB94688.1"
                     /db_xref="GI:21203990"
                     /translation="MAQDRKKVLVLGAGYAGLQTVTKLQKAISTEEAEITLINKNEYH
                     YEATWLHEASAGTLNYEDVLYPVESVLKKDKVNFVQAEVTKIDRDAKKVETNQGIYDF
                     DILVVALGFVSETFGIEGMKDHAFQIENVITARELSRHIEDKFANYAASKEKDDNDLS
                     ILVGGAGFTGVEFLGELTDRIPELCSKYGVDQNKVKITCVEAAPKMLPMFSEELVNHA
                     VSYLEDRGVEFKIATPIVACNEKGFVVEVDGEKQQLNAGTSVWAAGVRGSKLMEESFE
                     GVKRGRIVTKQDLTINGYDNIFVIGDCSAFIPAGEERPLPTTAQIAMQQGESVAKNIK
                     RILNGESTEEFEYVDRGTVCSLGSHDGVGMVFDKPIAGKKAAFMKKVIDTRAVFKIGG
                     IGLAFKKGKF"
     gene            1624. .3099
                     /gene="ampA"
     CDS             1624. .3099
                     /gene="ampA"
                     /note="ORFID:MW0824"
                     /codon_start=1
                     /transl_table=11
                     /product="probable cytosol aminopeptidase"
                     /protein_id="BAB94689.1"
                     /db_xref="GI:21203991"
                     /translation="MNFKLNNTLSNEINTLIIGIPEHLNQLERISFNHIDITESLERL
                     KHQHIIGSKVGKIYTTAFDVQDQTYRLITVGLGNLKTRSYQDMLKIWGHLFQYIKSEH
                     IEDTYLLMDSFISKYDQLSDVLMACGIQSERATYEFDHYKSSKKAPFKTNLNLISESL
                     IELDFIHEGISIGQSINLARDFSNMPPNVLTPQTFAEDIVNHFKNTKVKVDVKDYDTL
                     VSEGFGLLQAVGKGSKHKPRLVTITYNGKDKDEAPIALVGKGITYDSGGYSIKTKNGM
                     ATMKFDMCGAANVVGIIEAASRLQLPVNIVGVLACAENMINEASMKPDDVFTALSGET
                     VEVMNTDAEGRLVLADAVFYANQYQPSVIMDFATLTGAAIVALGDDKAAAFESNSKVI
                     LNDILQISSEVDEMVFELPITATERASIKHSDIADLVNHTNGQGKALFAASFVTHFSG
                     QTPHIHFDIAGPATTNKASYNGPKGPTGFMIPTIVQWLKQQ"
     gene            3510. .4826
                     /gene="MW0825"
     CDS             3510. .4826
                     /gene="MW0825"
                     /note="ORFID:MW0825"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB94690.1"
                     /db_xref="GI:21203992"
                     /translation="MINAVVIAVILMIVLCLCRLNVVISLFISALVGGLISGMSIEKV
                     INVFGKNIVDGAEVALSYALLGGFAALISYSGITDYLVGKIINAIHAENSRWSRVKVK
                     VTIIIALLAMSIMSQNLIPVHIAFIPIVIPPLLSLFNDLKIDRRLIGLIIGFGLCFPY
                     VLLPYGFGQIFQQIIQSGFAKANHPIEFNMIWKAMLIPSMGYIVGLLIGLYVYRKPRE
                     YETRKISDSDNVTELKPYILIVTIVAILATFLVQTFTDSMIFGALAGVLVFFISRAYN
                     WYELDAKFVEGIKIMAYIGVVILTANGFAGVMNATGDIDELVKTLTSITGDNKLFSII
                     MMYVIGLIVTLGIGSSFATIPIIASLFIPFGASIGLDTMALIALIGTASALGDSGSPA
                     SDSTLGPTAGLNVDGQHDHIRDTCVPNFLFYNIPLMIFGTIAAMVL"
     gene            4845. .5219
                     /gene="MW0826"
     CDS             4845. .5219
                     /gene="MW0826"
                     /note="ORFID:MW0826"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB94691.1"
                     /db_xref="GI:21203993"
                     /translation="MTHLLETFEMSIDHQEDGLVVISMPVTDKVKQPFGYLHGGASIA
                     LGETACSLGSANLIDTTKFIPLGLEMNANHIHSAKDGHVTATAEIIHQGKSTHVWDIK
                     IKNDKEQLITVMRGTVAIKPLK"
     gene            complement(5275. .6429)
                     /gene="MW0827"
     CDS             complement(5275. .6429)
                     /gene="MW0827"
                     /note="ORFID:MW0827"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAB94692.1"
                     /db_xref="GI:21203994"
                     /translation="MHWTIIGGGIQGTAIAQKLLSSGLTTDRLTIIDPHETFCQRFNS
                     FTNRIEMPYLRSPIVHHVHPQPFHLKQFAKQHQYTNAFYGPYQRPELTMFMDHIAHAS
                     KQYQLEDCLVQGLVQTLDKQEDKWHIKLEDGQIITTDCVVIAIGSTNIPFMPDILKDK
                     QNVNHIFEKEHDQVVYDKTDHIVGSGITAAHLALKLLNHDNDKKIHLWLNKDIEIHDF
                     DADPGWLGPKNMSTFLSTKSMPERNAIVQRERHKGSMPHELYLRLKKHIKNGRINVHK
                     TPITQISDGLINTENNSVPYQQIMVATGFEQDFMSQPLIKQLIQNYDAPINECNYPVI
                     SEKLEWIPNLFVAGCFADLELGPFGRNVMGGRKAAERIEQAFLKLQQYSA"
     gene            complement(6678. .7034)
                     /gene="mnhG"
     CDS             complement(6678. .7034)
                     /gene="mnhG"
                     /note="ORFID:MW0828"
                     /codon_start=1
                     /transl_table=11
                     /product="Na+/H+ antiporter subunit"
                     /protein_id="BAB94693.1"
                     /db_xref="GI:21203995"
                     /translation="MIKIILISLALIFVIIGALISALAAIGLLRLEDVYSRAHAAGKA
                     STLGAMSLLFGTFLYFIATQGFVNMQLIVAIIFVLITGPLSSHMIMKAAYNIKTPYTK
                     KTKVDEISEDLKDTKL"
     gene            complement(7012. .7305)
                     /gene="mnhF"
     CDS             complement(7012. .7305)
                     /gene="mnhF"
                     /note="ORFID:MW0829"
                     /codon_start=1
                     /transl_table=11
                     /product="Na+/H+ antiporter subunit"
                     /protein_id="BAB94694.1"
                     /db_xref="GI:21203996"
                     /translation="MNHNVIIVIALIIVVISMLAMLIRVVLGPSLADRVVALDAIGLQ
                     LMAVIALFSILLNIKYMIVVIMMIGILAFLGTAVFSKFMDKGKVIEHDQNHTD"
     gene            complement(7305. .7784)
                     /gene="mnhE"
     CDS             complement(7305. .7784)
                     /gene="mnhE"
                     /note="ORFID:MW0830"
                     /codon_start=1
                     /transl_table=11
                     /product="Na+/H+ antiporter subunit"
                     /protein_id="BAB94695.1"
                     /db_xref="GI:21203997"
                     /translation="MAVQLVLNFIIAVFWLFVTNSYTTNNFVLGFIFGLVLVYLLHRV
                     LPGRFYVITLYRIIKLVIIFLIELIKANFDVLKIIIKPSIKNEPGFFVYHTDLKKDWQ
                     IVLLSNLITLTPGTVVLGVSDDRTKIYIHAIDFSTKEQEVESIKTSLEKIVREVGEI"
     gene            complement(7786. .9282)
                     /gene="mnhD"
     CDS             complement(7786. .9282)
                     /gene="mnhD"
                     /note="ORFID:MW0831"
                     /codon_start=1
                     /transl_table=11
                     /product="Na+/H+-antiporter subunit"
                     /protein_id="BAB94696.1"
                     /db_xref="GI:21203998"
                     /translation="MIESNMLVLTLVIPVITAILLVFIGKRPIIKRYVALGGTLLTLV
                     AAIINLANVVKHGPIRVELGSWKAPYSIVFVLDIFSALLIITSIIITAIVILYSYQTI
                     GIERERYYYYFSVLFMLIGIIGAFTTGDIFNLFVFFEVFLMSSYFLLVIGSTKIQLQE
                     TIKYVLVNVVSSSFFVMGVAILYSVVGTLNLADISNKLANLSAHDSGLVNIVFILFIF
                     VFATKAGVFPMFVWLPSAYYAPPIPIIAFFGALLTKVGVYAIARTLSLFFSDNVSFSH
                     YVILFLALLTIIFGCVGAVAYANIKKIILYNVMIAVGVILVGVAMMTESGMIGAIYYT
                     LHDMLVKLALFLLIGIMIKITGTADLRQFGGLIKRYPVLGWSFFIAALSLAGIPPLSG
                     FYGKFFIVQSTFERGFYLSGVIVLLSSLVVLYSVIRIFLQGFFGQPKGYDLNNKVDVK
                     YLTTIAIVAVVITVLYGLSADYLYPMVKAGAETFYNPSTYVKAVLGGK"
     gene            complement(9275..9616)
                     /gene="mnhC"
     CDS             complement(9275..9616)
                     /gene="mnhC"
                     /note="ORFID:MW0832"
                     /codon_start=1
                     /transl_table=11
                     /product="Na+/H+ antiporter subunit"
                     /protein_id="BAB94697.1"
                     /db_xref="GI:21203999"
                     /translation="MEIIMIFVSGILTAISVYLVLSKSLIRIVMGTTLLTHAANLFLI
                     TMGGLKHGTVPIYEANVKSYVDPIPQALILTAIVIAFATTAFFLVLAFRTYKELGTDN
                     VESMKGVPEDD"
     gene            complement(9616..10044)
                     /gene="mnhB"
     CDS             complement(9616..10044)
                     /gene="mnhB"
                     /note="ORFID:MW0833"
                     /codon_start=1
                     /transl_table=11
                     /product="Na+/H+ antiporter subunit"
                     /protein_id="BAB94698.1"
                     /db_xref="GI:21204000"
                     /translation="MNRQQNDLILQFAAVIIFFMVMVFGFSLFLAGHYTPGGGFVGGL

  Query Match             26.1%;  Score 262.8;  DB 1;  Length 290150;
  Best Local Similarity   55.0%;  Pred. No. 2.8e-31;
  Matches  541;  Conservative    0;  Mismatches  437;  Indels    6;  Gaps    1;

Qy          17 GAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCAT 76
               ||||| ||    |    ||    || |    | |  |||| || ||   ||    ||
Db      107984 GAGGTATATGTGTGAATAGGAATATCGTTAAACTAGTTGTGTTTATGCTAATCTTAGTTG 108043

Qy          77 GTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCA 136
                 | |       |||   || |   |  ||  ||      ||  ||         ||| |
Db      108044 TAGCAGTAGCGGGTTGTGGTCAAAAAGATACTGAAGAGAAAACTGAAATGACGACAATAA 108103

Qy         137 AGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGT 196
               |  | |||  ||| |||   ||| |    ||| | ||||||||||||||||  | || |
Db      108104 AAGATGAATTAGGAACTGAAAAAATTAAGAAAAATCCTAAACGTGTTGTTGTATTAGAAT 108163

Qy         197 ATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
               ||   |||| ||||   ||||  || |||||| | ||||||||||| || || ||||||
Db      108164 ATAGTTTTGCTGATTATTTAGCAGCATTAGATATGAAACCTGTTGGTATTGCAGATGATG 108223

Qy         257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
                ||  |  ||   ||| |  ||  || |||||||||| ||||||   |||    || || |
Db      108224 GCAGCAGTAAAAATATAACAAAGTCAGTAAGAGATAAGATTGGGGCATATGAATCGGTTG 108283

Qy         317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
               || |  |   |||||| ||  | ||||  || |||||| | ||||| ||||| || ||||
Db      108284 GATCTAGACCGCAACCGAATATGGAAGTGATAAGTAAATTAAAACCGGATTTGATCATTG 108343

Qy         377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
               | |||  ||  ||||| ||    ||  |   ||| || |  ||||||||  || || |||
Db      108344 CAGATGTTAGCAGACATAAGAAAATCAAATCAGAATTGAGCAAAATTGCGCCGACAATTA 108403

Qy         437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
                  |    |||      |||||||| |||| ||||||||| || ||||||||| |  | |
Db      108404 TGTTAGTCAGTGGTACGGGAGATTACAATGCAAATATTGAAGCATTTAAAACAGTCGCTA 108463

Qy         497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
               ||||  | || |||||  |||||||  | || ||  | || | | || ||||| | | | |
Db      108464 AAGCAGTTGGCAAAGAGAAAGAAGGCGAGAAGCGTCTGGAAAAGCATGATAAAATATTAG 108523

Qy         557 AAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTG 616
                 ||   || |||  ||||      |  ||      ||   | |||  |    |
Db      108524 CGGAGATTAGAAAGAAAATTGAACAGAGTACGTTAAAATCTGCATTTGCATTCGGTATCT 108583

Qy         617 CTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAAC 676
               | | | ||||| || ||  | ||       |  | | | | || |||||| ||  | ||
Db      108584 CAAGAGCAGGTATGTTTATTAATAATGAAGATACATTTATGGGACAATTCTTACTTAAAA 108643

Qy         677 TAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAG 736
               | ||| || ||   | | | |   | ||    ||||    | |  || |          |
Db      108644 TGGGTATTCAACCTGAAGTCACAAAAGACAAAACTACGCATGTTGGTGAACGCAAGGGTG 108703

Qy         737 GACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCA 796
               | |||||  || |  | || | ||||    |  |  |  | |||||  |   |||| |
Db      108704 GTCCTTATATATATTTAAATAATGAAGAACTTGCCAATATCAATCCAAAAGTTATGATTT 108763

Qy         797 TAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTG 856
               ||   ||  ||     ||     ||  |         ||   || |      |   | | |
Db      108764 TAGCCACTGACGGGAAAACGGACAAAAA------TAGAACGAAATTCATTGATCCTGCAG 108817

Qy         857 TATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTAT 916
               | |||||   |||||| ||||||||| || |    ||| ||    | |||||  |    |
Db      108818 TTTGGAAATCATTAAAAGCTGTGAAAGATAACAAAGTTTATGACGTTGACCGAAATAAGT 108877

Qy         917 GGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTAT 976
               ||   | |||| | ||   |||  |    |||   |||||| |||| |  |   || |
Db      108878 GGTTGAAATCAAGGGGGATTATCGCAAGTGAAAGTATGGCAGAAGATTTAGAAAAAATTG 108937

Qy         977 CTAAGAAAGATAGTAAAAAAGATA 1000
               |  | ||||  |   ||||| | |
Db      108938 CAGAAAAAGCAAAATAAAAATACA 108961


RESULT 15
BX571856_10
WPCOMMENT
  Sequence split into 29 fragments  LOCUS BX571856 Accession BX571856
     Fragment Name        Begin       End
     BX571856_00              1    110000
     BX571856_01         100001    210000
     BX571856_02         200001    310000
     BX571856_03         300001    410000
     BX571856_04         400001    510000
     BX571856_05         500001    610000
     BX571856_06         600001    710000
     BX571856_07         700001    810000
     BX571856_08         800001    910000
     BX571856_09         900001   1010000
     BX571856_10        1000001   1110000
     BX571856_11        1100001   1210000
     BX571856_12        1200001   1310000
     BX571856_13        1300001   1410000
     BX571856_14        1400001   1510000
     BX571856_15        1500001   1610000
     BX571856_16        1600001   1710000
     BX571856_17        1700001   1810000
     BX571856_18        1800001   1910000
     BX571856_19        1900001   2010000
     BX571856_20        2000001   2110000
     BX571856_21        2100001   2210000
     BX571856_22        2200001   2310000
     BX571856_23        2300001   2410000
     BX571856_24        2400001   2510000
     BX571856_25        2500001   2610000
     BX571856_26        2600001   2710000
     BX571856_27        2700001   2810000
     BX571856_28        2800001   2902619
 Continuation (11 of 29) of BX571856 from base 1000001 (BX571856 Staphylococcus aureus su

  Query Match             25.2%;  Score 254.2;  DB 1;  Length 110000;
  Best Local Similarity   54.3%;  Pred. No. 7.2e-30;
  Matches  539;  Conservative    0;  Mismatches  448;  Indels    6;  Gaps    1;

Qy           8 AATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAA 67
               |||   |  |||||  |    |    ||    || |    |||  |||| || ||   ||
Db       54194 AATTGTTTGGAGGTAAATGTGTGAATAGGAATATCGTTAAATTAGTTGTGTTTATGCTAA 54253

Qy          68 CTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAG 127
                   ||     | |       |||   || |   |  ||   ||      ||  ||
Db      54254 TCTTAGTTGTAGCAGTAGCGGGTTGTGGTCAAAAAGATACTGAAGAGAAAACTGAAATGA 54313

Qy        128 TTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTG 187
                   || ||  | |||  ||| |||   ||| |    ||| | || |||||| ||||||
Db      54314 CGACGATAAAAGATGAATTAGGAACTGAAAAAATTAAGAAAAATCCGAAACGTATTGTTG 54373

Qy        188 TTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAG 247
              |  | || ||   |||| ||||   ||||  || ||||| | ||||||||||| || |
Db      54374 TATTAGAATATAGTTTTGCTGATTATTTAGCAGCATTAGATATGAAACCTGTTGGTATTG 54433

Qy        248 CGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACA 307
              | ||||||   |  |  ||   ||| |  ||  || ||||||||||| |||||   |||
Db      54434 CAGATGATGGTAGTACTAAAAATATAACAAAGTCAGTAAGAGATAAGATTGGGGCATATG 54493

Qy        308 CTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATT 367
                || || ||| |  |   |||||| ||  | ||||  || |||||| | ||||| ||||
Db      54494 AATCGGTTGGATCTAGACCGCAACCGAATATGGAAGTGATAAGTAAATTAAAACCGGATT 54553

Qy        368 TAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTC 427
              | |||||||| |||   |  ||||| ||    ||  |   ||| || |  |||||||| |
Db      54554 TGATTATTGCAGATGTCAGCAGACATAAGAAAATCAAATCAGAATTGAGCAAAATTGCGC 54613

Qy        428 CTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAA 487
              | || ||     |    ||       ||||||||| |||| ||||||||| || |||||||
Db      54614 CGACAATCATGTTAGTTAGCGGTACGGGAGATTACAATGCAAATATTGAAGCATTTAAAA 54673

Qy        488 CAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATA 547
              || |  | ||||||  | || |||||  |||||||  ||||  ||  | ||| | ||  |||
Db      54674 CAGTCGCTAAAGCAGTTGGCAAAGAGAAAGAAGGCGAAAAGCGTCTGGAAAAGCATAATA 54733

Qy        548 AGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTG 607
              | | | | |  ||   || |||  ||||      |  ||      ||   | |||  |
Db      54734 AAATATTAGCGGAGATTAGAAAGAAAATTGAACAGAGTACGTTAAAATCTGCATTTGCAT 54793

Qy        608 CAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCC 667
                |      | | | ||||| || ||  | ||       |    | |   || ||||||
Db      54794 TTGGTATCTCAAGAGCAGGTATGTTTATTAATAATGAAGATACATTTATGGGACAATTCT 54853

Qy        668 TAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGT 727
              ||| | || | ||| || ||   | | | |   | ||    |||     | |  || |
Db      54854 TAATTAAAATGGGTATTCAACCTGAAGTCACAAAAGACAAAACTGCGCATGTTGGTGAAC 54913

Qy        728 ATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGC 787
                      || |||||  |  |  | || | ||||    |  |  |  |||||  |
Db      54914 GCAAGGGTGGCCCTTATATTTATTTAAATAATGAAGAGCTTGCCAATATCAATCCAAAAG 54973

Qy        788 GTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAA 847
               |||| |  ||   || ||   |  ||    |||  |         ||   || |
Db      54974 TTATGATTTTAGCTACGAATGGAAAAACAGATAAAAA------TAGAACGAAATTCATTG 55027

Qy        848 AAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACC 907
              |   | | || |||||   |||||| ||||||||||| || |    ||  ||  | | || |
Db      55028 ATCCTGCAGTTTGGAAATCATTAAAAGCTGTGAAAGATAATAAAGTATATGATGTTGATC 55087

Qy        908 GTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTG 967
              |  |    |||     |||| | ||| | ||  |    |||   ||||||| ||||  | |
Db      55088 GAAATAAGTGGTTGCAATCAAGAGGTATTATGGCAAGTGAAAGTATGGCAGAAGATTTAG 55147

Qy        968 TTGAATTATCTAAGAAAGATAGTAAAAAAGATA 1000
                 || |  |  | ||||  |  |||||| | |
Db      55148 AAAAAATTGCAGAAAAAGCAAAATAAAAATACA 55180

RESULT 16
AR354237
LOCUS       AR354237                 668 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 355 from patent US 6593114.
ACCESSION   AR354237
VERSION     AR354237.1  GI:33760321
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
```

```
REFERENCE   1  (bases 1 to 668)
  AUTHORS   Kunsch,C.A., Choi,G.H., Barash,S., Dillon,P.J., Fannon,M.R. and
            Rosen,C.A.
  TITLE     Staphylococcus aureus polynucleotides and sequences
  JOURNAL   Patent: US 6593114-A 355 15-JUL-2003;
FEATURES             Location/Qualifiers
     source          1. .668
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             22.1%;  Score 222.6;  DB 6;  Length 668;
  Best Local Similarity   74.8%;  Pred. No. 1.7e-24;
  Matches  279;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy        627 TTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAA 686
              |||  | || ||||||| | | || ||||||||| |||||  |||   |||||| |||||
Db          1 TTTATTAGCACATCCAAACTATTCATATGTTGGACAATTTTTAAACGAACTAGGATTTAA 60

Qy        687 AGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTT 746
              | | ||||||||||| ||||| || ||||||||||||||| ||| | ||||||||||||||
Db         61 AAATGCATTAAGTGACGATGTAACAAAAGGTTTAAGTAAATATTTGAAAGGACCTTACTT 120

Qy        747 ACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAA 806
              |||| |  ||||||||  ||||  | |  | ||||| |||||||||| ||||| |||||| |
Db        121 ACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGATCATTATGACAGA 180

Qy        807 CAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAA 866
                | || |      |     ||  | | ||  |  || || ||||| |   |||||| ||
Db        181 TCATGCTAAAAAAGATTCTGCTGAATTCAAGAAGTTACAAGAAGATGCAACATGGAAAAA 240

Qy        867 ATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATC 926
               || || || || |||||| | || || |||| | ||||||||  | |||||||||||
Db        241 GTTGAATGCAGTTAAAAATAATCGCGTGGATATTGTTGACCGTGATGTTTGGGCAAGATC 300

Qy        927 ACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGA 986
               ||||| ||||||||||| ||||||||||| ||||||||||||||||||||| || |||||
Db        301 TCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTTGTTGAATTATCAAAAAAAGA 360

Qy        987 TAGTAAAAAAGAT 999
                  ||  ||| |
Db        361 ACAAAAGTAAGGT 373

RESULT 17
AR535793
LOCUS       AR535793                 668 bp    DNA     linear   PAT 08-OCT-2004
DEFINITION  Sequence 355 from patent US 6737248.
ACCESSION   AR535793
VERSION     AR535793.1  GI:53927010
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 668)
  AUTHORS   Kunsch,C.A., Choi,G.A., Barash,S.C., Dillon,P.J., Fannon,M.R. and
            Rosen,C.A.
  TITLE     Staphylococcus aureus polynucleotides and sequences
  JOURNAL   Patent: US 6737248-A 355 18-MAY-2004;
FEATURES             Location/Qualifiers
     source          1. .668
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             22.1%;  Score 222.6;  DB 6;  Length 668;
  Best Local Similarity   74.8%;  Pred. No. 1.7e-24;
  Matches  279;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy        627 TTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAA 686
              |||  | || ||||||| | | || ||||||||| |||||  |||   |||||| |||||
Db          1 TTTATTAGCACATCCAAACTATTCATATGTTGGACAATTTTTAAACGAACTAGGATTTAA 60
```

```
Qy        687 AGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTT 746
              | | ||||||||||| ||||| || |||||||||||||| ||| | ||||||||||||||
Db         61 AAATGCATTAAGTGACGATGTAACAAAAGGTTTAAGTAAATATTTGAAAGGACCTTACTT 120

Qy        747 ACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAA 806
              |||| |  ||||||||  |||| || |  | ||||| ||||||||| |||| |||||| |
Db        121 ACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGATCATTATGACAGA 180

Qy        807 CAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAA 866
                | || |      |     ||  | | ||  |  || || ||||| |   |||||| ||
Db        181 TCATGCTAAAAAAGATTCTGCTGAATTCAAGAAGTTACAAGAAGATGCAACATGGAAAAA 240

Qy        867 ATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATC 926
               || || || || |||||| | || || |||||| | ||||||||  | |||||||||||
Db        241 GTTGAATGCAGTTAAAAATAATCGCGTGGATATTGTTGACCGTGATGTTTGGGCAAGATC 300

Qy        927 ACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGA 986
               ||||| ||||||||||| ||||||||||| |||||||||||||||||||| || |||||
Db        301 TCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTTGTTGAATTATCAAAAAAAGA 360

Qy        987 TAGTAAAAAAGAT 999
                  ||  ||| |
Db        361 ACAAAAGTAAGGT 373


RESULT 18
AX951891/c
LOCUS       AX951891                5718 bp    DNA     linear   PAT 08-JAN-2004
DEFINITION  Sequence 20 from Patent WO03095658.
ACCESSION   AX951891
VERSION     AX951891.1  GI:40782280
KEYWORDS    .
SOURCE      Bacillus subtilis
  ORGANISM  Bacillus subtilis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE   1
  AUTHORS   bj Rnvad,M.E., j Rgensen,P.L. and Hansen,P.K.
  TITLE     Homologous recombination into bacterium for the generation of
            polynucleotide libraries
  JOURNAL   Patent: WO 03095658-A 20 20-NOV-2003;
            Novozymes A/S (DK)
FEATURES             Location/Qualifiers
     source          1..5718
                     /organism="Bacillus subtilis"
                     /mol_type="unassigned DNA"
                     /strain="MB1510"
                     /db_xref="taxon:1423"
ORIGIN

  Query Match             18.9%;  Score 190.8;  DB 6;  Length 5718;
  Best Local Similarity   55.0%;  Pred. No. 6.5e-20;
  Matches  488;  Conservative    0;  Mismatches  357;  Indels   43;  Gaps    4;

Qy        115 TCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCT 174
              || || |||      |  ||| |  |||||||||| |  |||| ||||         |||
Db        931 TCTAAGGATTCCAGAGTGATCCATGACGAAGAAGGAAAAACGACAGTAAGCGGCACACCT 872

Qy        175 AAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAA 234
              || || || ||||| ||||||  ||||||  | |||||  |      | | |   |||
Db        871 AAGCGGGTGGTTGTGCTTGAGCTTTCATTCTTGGATGCCGTTCACAATCTCGGCATTACG 812

Qy        235 CCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAA 294
              || || || || || ||||| ||||||||| ||   ||||  || |   |  |
Db        811 CCGGTGGGCATCGCAGATGACAACAAAAAAGATATGATTAAAAAGCTTGTCGGCAGCTCC 752

Qy        295 ATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAA 354
              ||||   | ||||| |||||||| ||||| |   |||| ||  | || |  ||||||
Db        751 ATTG---ATTACACATCTGTAGGCACACGCAGCGAACCCAATCTTGAGGTCATCAGTTCC 695

Qy        355 CTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTA 414
               | || || |||||||| || |||||   | |  | || |||   ||||||||| |  |
Db        694 TTGAAGCCTGATTTAATCATCGCTGACGCTGAGCGCCATAAAAACATTTATAAACAGCTG 635
```

```
Qy        415 AATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATT 474
              || ||||| || || ||||||||| | |||||    || |     ||| | ||||  ||
Db        634 AAAAAAATCGCCCCGACGATTGAATTAAAAAGCCGTGAAGCGACATATGACGAAACGATC 575

Qy        475 GATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTA 534
              ||    ||||  || ||| | ||||| |||  ||||||||| ||||| ||| ||    |
Db        574 GACAGCTTTACGACCATTGCTAAAGCATTAAATAAAGAAGATGAAGGAAAAGAAAAGCTT 515

Qy        535 GAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAA 594
              |  || ||| | ||    ||  | ||    |||   ||| |  | |  ||| |||| |
Db        514 GCCGAGCACAAAAAAGTCATCAACGATCTAAAAGCCGAACTTCCGAAAGATGAAAACCGC 455

Qy        595 AAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTAT 654
              ||  |  | |  |  || ||             || |    |||| ||      || |||
Db        454 AACATCGTTCTCGGCGTTGCAAGAGCGGATTCCTTCCAGCTTCATACATCATCATCCTAT 395

Qy        655 GTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAA 714
              |  ||  || || | |     ||||| |||| | | ||  | |     |||
Db        394 GACGGAGAAATCTTTAAAATGCTAGGCTTTACACACGCTGTGAAGTCAGAT--------- 344

Qy        715 GGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAA 774
                                       | ||   | |  | | |  |||    ||     ||
Db        343 ---------------------AACGCCTATCAAGAGGTCAGCCTTGAGCAATTGAGCAAA 305

Qy        775 GTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
               |  |||||||   | |||||||     | ||| |||  |   | |  |         ||
Db        304 ATCGATCCTGATATTTTGTTCATCTCAGCCAACGAAGGCAAAACCATTG---------TA 254

Qy        835 AAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTT 894
               | ||     |||   | ||  | ||||| ||  | || || |||||||||  ||  ||
Db        253 GATGAGTGGAAAACGAACCCGCTCTGGAAAAATCTCAAAGCGGTGAAAAATGGACAAGTC 194

Qy        895 GATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATG 954
               ||  |   |||||||||   ||| ||||||   | ||  | |  |||   |||  ||
Db        193 TATGATGCGGACCGTGACACTTGGACAAGATTCAGAGGCATCAAGTCTAGTGAAACAAGC 134

Qy        955 GCAAAAGAACTTGTTGAATTA-TCTAAGAAAGATAGTAAAAAAGATAA 1001
              || |||||  |  || ||  | | ||  | | |||| | |||  || ||
Db        133 GCCAAAGATGTGCTTAAAAAAGTGTATAATAAATAGGACAAATGAGAA 86


RESULT 19
D86417
LOCUS       D86417                 37900 bp    DNA     linear   BCT 18-SEP-1997
DEFINITION  Bacillus subtilis 35.7 kb genomic DNA, 70-73 degree region,
            complete cds.
ACCESSION   D86417
VERSION     D86417.1  GI:2443217
KEYWORDS    YfmN; YfmL; YfmK; YfmH; YfmG; YfmA; YflT; Pel; YflS; YflR; YflQ;
            YflP; YflO; YflN; YflM; YflK; YflF; YflD; YflC; YflB; YflA; YfmM;
            YfmJ; YfmI; YfmF; YfmE; YfmD; YfmC; YfmB; YflL; YflJ; YflI; YflH;
            YflG; YflE; YfkT; YfkS; YfkR; YfkQ; TreP.
SOURCE      Bacillus subtilis
  ORGANISM  Bacillus subtilis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE   1  (sites)
  AUTHORS   Yamamoto,H., Uchiyama,S., Nugroho,F.A. and Sekiguchi,J.
  TITLE     Cloning and sequencing of a 35.7 kb in the 70 degree-73 degree
            region of the Bacillus subtilis genome reveal genes for a new
            two-component system, three spore germination proteins, an iron
            uptake system and a general stress response protein
  JOURNAL   Gene 194 (2), 191-199 (1997)
  MEDLINE   97417488
   PUBMED   9272861
REFERENCE   2  (bases 1 to 37900)
  AUTHORS   Sekiguchi,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-1996) Junichi Sekiguchi, Textile Science and
            Technology, Shinshu University, Department of Applied Biology;
            3-15-1 Tokida, Ueda-shi, Nagano 386, Japan
            (E-mail:jsekigu@giptc.shinshu-u.ac.jp, Tel:0268-21-5344,
            Fax:0268-21-5331)
FEATURES             Location/Qualifiers
```

```
source          1. .37900
                /organism="Bacillus subtilis"
                /mol_type="genomic DNA"
                /strain="AC327"
                /db_xref="taxon:1423"
CDS             complement(56. .1468)
                /codon_start=1
                /transl_table=11
                /product="TreP"
                /protein_id="BAA22289.1"
                /db_xref="GI:2443218"
                /translation="MGELNKSARQIVEAVGGAENIAAATHCVTRLRFALIDESKVDQE
                MLDQIDVVKGSFSTNGQFQVVIGQGTVNKVYAELVKETGIGESTKDEVKKASEKNMNP
                LQRAVKTLADIFIPILPAIVTAGLLMGINNILTAEGIFFSTKSIVQVYPQWADLANMI
                NLIAGTAFTFLPALIGWSAVKRFGGNPLLGIVLGVMLVHPDLLNAWGYGAAEQSGEIP
                VWNLFGLEVQKVGYQGQVLPILLASYMLAKIEVFLTKRTPEGIQLLVVAPITLLLTGF
                ASFIIIGPITFAIGNVLTSGLISVFGSFAALGGLLYGGFYSALVITGMHHTFLAVDLQ
                LIGSKLGGTFLWPMLALSNIAQGSAALAMMFIVKDEKQKGLSLTSGISAYLGITEPAI
                FGVNLRYRFPFIIAMVSSGLAGMYISSQGVLASSVGVGGVPGIFSIMSQYWGAFAIGM
                AIVLIVPFAGTYAYARFKHK"
RBS             complement(1466. .1483)
RBS             1643. .1663
CDS             1661. .3202
                /codon_start=1
                /transl_table=11
                /product="YfkQ"
                /protein_id="BAA22290.1"
                /db_xref="GI:2443219"
                /translation="MNKRDKALQLIKELEDNQDRPISPNLKENLENLTLLTEGCSDIV
                FRQFDFGNGLCGFIVYIEGIVKSEHIQDHALRPFLMHLTDQIDEHEEALQNTLSISSV
                ALETSMSKVAASIIEGNAVLFADGHSKGLILNIKGGQRRSIEEPITESTIRGSREGFT
                ESLRVNTALVRFRVKTFQLKMISFKIGTKTKTDVVLAYIDGLADPKVIDKAKKRIKKI
                KIDAVLESGYIEEFIEDDTYSPFPQLQYTERPDTVAAQLLEGRFAIFTDNTPFVLTGP
                ITFWQLMQASEDYYERYLMSNLIRWLRYMFLFVALYLPAIYVAVITYHQDLMPTNLMF
                SVASAREPIPFPAIIEALIMEISFEALREAGVRLPKTIGQTVSILGALVIGTAAVEAG
                IVSAPMVIIVSLTGIASFTIPRFNLAISIRMLRFPLMFLASIFGIFGIMLGTIILVLH
                LCKLQSFGIPYLSGISPFKRDEVKDIFVRAPWWTMTRRPGTYSRGNGQKGAKREDPKD
                EENNI"
RBS             3166. .3185
CDS             3183. .4337
                /codon_start=1
                /transl_table=11
                /product="YfkR"
                /protein_id="BAA22291.1"
                /db_xref="GI:2443220"
                /translation="MKKTIYKCVLPLLICILLTGCWDRTEINDIAFVVSSAIDKKKDQ
                YRVAMQIPLVGQLGGQTGGGGGTAGSKTWYVDSASGTTIREANNKLQTSLSRTINTSH
                RRTVIIGEDMARDGVAPVFDILTRNPQNRLTALILVSRGEARDILNTDVQLEQFPAEM
                IRELAIIATSRPVFLSRFMSDLVEIGSDPYAPVISASKTKPGGKGKSNLKIDGLAIFK
                KDRLMDIFKDEHMTAALMLLNQARQPEFIVDLPNQMGQASIQLQKSNASFHAAEKNGK
                LSMTIEIRAKGIIMENQSTYETRENDQYYIIQKALNRTIKKDVISTVHRLQKLKADPA
                GFQDRTIRSTATTKNLLKKEWEEVYKDMEVHVVPIVTIEQGGVLYKTISH"
RBS             4336. .4355
CDS             4353. .4553
                /codon_start=1
                /transl_table=11
                /product="YfkS"
                /protein_id="BAA22292.1"
                /db_xref="GI:2443221"
                /translation="MISYIVQTLIVCIAIYAYEWKNFRSANNLTKWAFSLLIAGSAFL
                WIYMRVNPLLPRLGHLFKYIPF"
RBS             4563. .4579
CDS             4577. .5653
                /codon_start=1
                /transl_table=11
                /product="YfkT"
                /protein_id="BAA22293.1"
                /db_xref="GI:2443222"
                /translation="MDKTSAYQGLFFGALYTLAVGLKHAPILMIESAKQNAWHSYILG
                VVIVIPALWLMHRLMKKHQDKNIYELLSDSSPIAGRIIILLFSLYFLLINAHDIRFFI
                NLINILFLPRTPMAVLGGVIIFVAICIAREGKETLTRMAQIFLFPFGILVLFLPFTLA
                TQIELQNLTPVFEGLIPYLQSGYYAFGTMGELIILPLLFSNRSVPLKYTIFAILLGAL
                LLAVMLFSSISVFGPNLTSTFFDPAYMVIRQIRITDFLDRSDLIIAAFWIPVIMVKIA
```

```
                GSLYIVVYGLSFLHSKIDPKAMYTPTGMFSVVCGFWFFLNTNQLIDFNRIKPIINVVI
                SLLLPLLIYLIIKSKALFGAKAKH"
CDS             complement(5650. .7065)
                /codon_start=1
                /transl_table=11
                /product="YflA"
                /protein_id="BAA22294.1"
                /db_xref="GI:2443223"
                /translation="MERLLVWIEHISDWLWGPPLIILLTGTGLYFTILLKGFQFRYPL
                YIFKQTIGSVGKKPKGEGTVTPLQALTSALSSTIGAANIVGVPAAIMFGGPGAVFWMW
                LIALFAMAIKFSESVLAVHYREKNEQGEYVGGPMYYITKGLRMKWLGVFFSVALIVEL
                IPSIMVQGNSVSVSLAETFSFNKIYAGIGIAFLIGLVVIGGVKRIGKVTEFVVPLMAG
                AYAGAGLLIVLMNLSSVPAFFSLVFSNAFTSSSAVGGFAGAALAETVRWGFARGLYSN
                EAGMGTAPIAHAAAMTDHPVRQGFWSVIGIVIDTLIICTTTAFIVLASGVWTGKNASN
                DPAALTTAAFQHYFGSGGGYFVSVSLVFFVVSTIMVVIFYGVKQAEFLFGRLAGHVIK
                FVYLAAIIIGAAGGAKAIWGVLDLALVFIVVPNVIALLLLSRKVKALYTEFFTSEQYY
                LKDIRKTKQKPVYPTKEAKNS"
RBS             complement(7063. .7077)
terminator      complement(7149. .7175)
CDS             complement(7190. .7450)
                /codon_start=1
                /transl_table=11
                /product="YflB"
                /protein_id="BAA22295.1"
                /db_xref="GI:2443224"
                /translation="MGYRILKNAGMAEDEGALKKELMTIDHLIEKCYDEKEREQLIRK
                KSEKQLLLDKLVDKKGMFSKPASAFYKNKVYDRLGRNRPSSS"
CDS             complement(7432. .7584)
                /codon_start=1
                /transl_table=11
                /product="YflC"
                /protein_id="BAA22296.1"
                /db_xref="GI:2443225"
                /translation="MMDFSHIVSEDKIKRAIKDGDFQNLPGMGKPLPKMTRHTCRNRC
                AWAIVF"
RBS             complement(7448. .7466)
RBS             complement(7582. .7596)
CDS             complement(7600. .7737)
                /codon_start=1
                /transl_table=11
                /product="YflD"
                /protein_id="BAA22297.1"
                /db_xref="GI:2443226"
                /translation="MREKVAKNAVESTFRFDITKCKTRYLSRNKGIKWYIENCMIKYK
                V"
RBS             complement(7735. .7752)
RBS             7824. .7840
CDS             7838. .9787
                /codon_start=1
                /transl_table=11
                /product="YflE"
                /protein_id="BAA22298.1"
                /db_xref="GI:2443227"
                /translation="MKTFIKERGLAFFLIAVVLLWIKTYVGYVLNFNLGIDNTIQKIL
                LFVNPLSSSLFFLGFGLLFKKKLQQTAIIVIHFLMSFLLYANIVYYRFFNDFITIPVI
                MQAKTNGGQLGDSAFSLMRPTDAFYFIDTIILIIILAIKVNKPAETSSKKSFRIIFASS
                ILVFLINLAVAESDRPELLTRSFDRNYLVKYLGTYNFTIYDAVQNIKSNSQRALADSS
                DVTEVENYMKANYDVPNNVYFGKAEGKNVIYVSLESLQSFIIDYKIDGKEVTPFLNKL
                AHDNETFYFDNFFHQTGQGKTSDAEFMMENSLYPLAQGSVFVNKAQNTLQSVPAILKS
                KNYTSATFHGNTQTFWNRNEMYKAEGIDKFFDSAYYDMNEENTKNYGMKDKPFFKESM
                PLLESLPQPFYTKFITLSNHFPFGMDEGDTDFPAGDFGDSVVDNYFQSAHYLDQSIEQ
                FFNDLKKDGLYDKSIIVMYGDHYGISENHNKAMAKVLGKDEITDYDNAQLQRVPLFIH
                AAGVKGEKVHKYAGDVDVAPTILHLLGVDTKDYLMSGSDILSKEHREVIPFRNGDFIS
                PKYTKISGKYYDTKTGKELDESEVDKSEDSLVKKELEMSDKIINGDLLRFYEPKGFKK
                VNPSDYDYTKHDEDSSETSKDNEDK"
terminator      complement(9789. .9820)
CDS             complement(9820. .11178)
                /codon_start=1
                /transl_table=11
                /product="YflF"
                /protein_id="BAA22299.1"
                /db_xref="GI:2443228"
                /translation="MLSFLQKLGKSFMLPIAVLPAVGIILALGREDVFNIPFVYQAGT
```

```
                    AVFDHLPLIFAIGIAIGISKDSNGAAGLSGAISYLMLDAATKTIDKTNNMAVFGGIIA
                    GLIAGYTYNRFKDTKLPEYLGFFSGRRLVPILTAIITIILAGIFGVVWPPIQSCINSF
                    GEWMLGLGGIGAGIFGLFNRLLIPLGLHHVLNNIFWFQFGEYNGVTGDLARFFAKDPT
                    AGTYMTGFFPIMMFGLPAACLAMVVTAKPSKRKATAGMMIGFALTAFITGITEPIEFA
                    FMFLSPLLYAVHAVLTGLSLFIVNWLGIRSGFSFSAGAIDYVLSYGIAEKPLLLLLVG
                    ICYAAVYFIVFYVLIKALNLKTPGREDDDVDEVLDENTVQDVNENIMLKGLGGKENLQ
                    TIDHCATRLRLTVKDTALVDEALLKKAGAKGVVKSGGQSVQVIIGPNVEFAAEELRAA
                    VK"
     RBS            complement(11176. .11195)
     RBS            11334. .11352
     CDS            11350. .12099
                    /codon_start=1
                    /transl_table=11
                    /product="YflG"
                    /protein_id="BAA22300.1"
                    /db_xref="GI:2443229"
                    /translation="MIVTNDQELEGLKKIGRIVALAREEMKRKAEPGMSTKDLDLIGK
                    AVLDEHGAVSAPEKEYDFPGVTCISVNDEVAHGIPSTSKILKAGDLVNIDISAEFGGF
                    YSDTGISFVLGEGEERLHKLCQCAENAFQKGLQQAKAGKRQNQIGRAVYHEARSQGFT
                    VIKTLTGHGIGRSLHEAPNHIMNYYDPFDNALFKNGTVIALEPFISTKAETIVEAGDG
                    WTFKTPDKSMVAQVEHTIVITKDEPIILTKL"
     RBS            12166. .12183
     CDS            12181. .12495
                    /codon_start=1

  Query Match             18.9%;  Score 190.8;  DB 1;  Length 37900;
  Best Local Similarity   55.0%;  Pred. No. 4e-20;
  Matches  488;  Conservative    0;  Mismatches  357;  Indels   43;  Gaps    4;

Qy        115 TCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCT 174
              || || |||      |  ||| |  |||||||||| |  |||| ||||          |||
Db      25213 TCTAAGGATTCCAGAGTGATCCATGACGAAGAAGGAAAAACGACAGTAAGCGGCACACCT 25272

Qy        175 AAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAA 234
              || || || ||||| |||||||  ||||||  | |||||  |      | | |   |||
Db      25273 AAGCGGGTGGTTGTGCTTGAGCTTTCATTCTTGGATGCCGTTCACAATCTCGGCATTACG 25332

Qy        235 CCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAA 294
              || || || || || ||||| ||||||||| ||   ||||  || | | |  |
Db      25333 CCGGTGGGCATCGCAGATGACAACAAAAAAGATATGATTAAAAAGCTTGTCGGCAGCTCC 25392

Qy        295 ATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAA 354
              ||||   | ||||| |||||||| ||||| |   |||| ||  | || |  ||||||
Db      25393 ATTG---ATTACACATCTGTAGGCACACGCAGCGAACCCAATCTTGAGGTCATCAGTTCC 25449

Qy        355 CTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTA 414
               | || || |||||||| || |||||   | |  | || |||   |||||||||| |  |
Db      25450 TTGAAGCCTGATTTAATCATCGCTGACGCTGAGCGCCATAAAAACATTTATAAACAGCTG 25509

Qy        415 AATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATT 474
              || ||||| || || ||||||||||| | |||||    || |     ||| | ||||  ||
Db      25510 AAAAAAATCGCCCCGACGATTGAATTAAAAAGCCGTGAAGCGACATATGACGAAACGATC 25569

Qy        475 GATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTA 534
              ||    ||||  || ||| | ||||| |||  ||||||||| ||||| ||| ||    |
Db      25570 GACAGCTTTACGACCATTGCTAAAGCATTAAATAAAGAAGATGAAGGAAAAGAAAAGCTT 25629

Qy        535 GAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAA 594
              |  || ||| | ||    ||  | ||    |||   ||| |  | |  ||| | |||| |
Db      25630 GCCGAGCACAAAAAAGTCATCAACGATCTAAAAGCCGAACTTCCGAAAGATGAAAACCGC 25689

Qy        595 AAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTAT 654
              ||  |  | |  |  || ||             || |    |||| ||      || |||
Db      25690 AACATCGTTCTCGGCGTTGCAAGAGCGGATTCCTTCCAGCTTCATACATCATCATCCTAT 25749

Qy        655 GTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAA 714
              |  ||  || || | |     ||||| |||| | | ||  | |     |||
Db      25750 GACGGAGAAATCTTTAAAATGCTAGGCTTTACACACGCTGTGAAGTCAGAT--------- 25800

Qy        715 GGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAA 774
                                       | ||   | |  | | |  |||    ||     ||
Db      25801 ---------------------AACGCCTATCAAGAGGTCAGCCTTGAGCAATTGAGCAAA 25839

Qy        775 GTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
               |  |||||||   | |||||||     | ||| ||   |   | |  |         ||
Db      25840 ATCGATCCTGATATTTTGTTCATCTCAGCCAACGAAGGCAAAACCATTG---------TA 25890

Qy        835 AAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTT 894
               | ||     |||   | ||  | ||||| ||  | || || ||||||||||  ||  ||
Db      25891 GATGAGTGGAAAACGAACCCGCTCTGGAAAAATCTCAAAGCGGTGAAAAATGGACAAGTC 25950

Qy        895 GATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATG 954
               ||  |   |||||||||   ||| |||||||   | ||  | |  |||    |||  ||
Db      25951 TATGATGCGGACCGTGACACTTGGACAAGATTCAGAGGCATCAAGTCTAGTGAAACAAGC 26010

Qy        955 GCAAAAGAACTTGTTGAATTA-TCTAAGAAAGATAGTAAAAAAGATAA 1001
              || |||||  |  || ||  | | ||  | | |||| | ||| || ||
Db      26011 GCCAAAGATGTGCTTAAAAAAGTGTATAATAAATAGGACAAATGAGAA 26058


RESULT 20
BSUB0005/c
LOCUS       BSUB0005               197409 bp    DNA     linear   BCT 07-JUL-2003
DEFINITION  Bacillus subtilis complete genome (section 5 of 21): from 813670 to
            1011078.
ACCESSION   Z99108 AL009126
VERSION     Z99108.2  GI:32468715
KEYWORDS    .
SOURCE      Bacillus subtilis subsp. subtilis str. 168
  ORGANISM  Bacillus subtilis subsp. subtilis str. 168
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE   1  (bases 1 to 197409)
  AUTHORS   Kunst,F., Ogasawara,N., Moszer,I., Albertini,A.M., Alloni,G.,
            Azevedo,V., Bertero,M.G., Bessieres,P., Bolotin,A., Borchert,S.,
            Borriss,R., Boursier,L., Brans,A., Braun,M., Brignell,S.C.,
            Bron,S., Brouillet,S., Bruschi,C.V., Caldwell,B., Capuano,V.,
            Carter,N.M., Choi,S.K., Codani,J.J., Connerton,I.F., Cummings,N.J.,
            Daniel,R.A., Denizot,F., Devine,K.M., Dusterhoft,A., Ehrlich,S.D.,
            Emmerson,P.T., Entian,K.D., Errington,J., Fabret,C., Ferrari,E.,
            Foulger,D., Fritz,C., Fujita,M., Fujita,Y., Fuma,S., Galizzi,A.,
            Galleron,N., Ghim,S.Y., Glaser,P., Goffeau,A., Golightly,E.J.,
            Grandi,G., Guiseppi,G., Guy,B.J., Haga,K., Haiech,J., Harwood,C.R.,
            Henaut,A., Hilbert,H., Holsappel,S., Hosono,S., Hullo,M.F.,
            Itaya,M., Jones,L., Joris,B., Karamata,D., Kasahara,Y.,
            Klaerr-Blanchard,M., Klein,C., Kobayashi,Y., Koetter,P.,
            Koningstein,G., Krogh,S., Kumano,M., Kurita,K., Lapidus,A.,
            Lardinois,S., Lauber,J., Lazarevic,V., Lee,S.M., Levine,A., Liu,H.,
            Masuda,S., Mauel,C., Medigue,C., Medina,N., Mellado,R.P.,
            Mizuno,M., Moestl,D., Nakai,S., Noback,M., Noone,D., O'Reilly,M.,
            Ogawa,K., Ogiwara,A., Oudega,B., Park,S.H., Parro,V., Pohl,T.M.,
            Portetelle,D., Porwollik,S., Prescott,A.M., Presecan,E., Pujic,P.,
            Purnelle,B., Rapoport,G., Rey,M., Reynolds,S., Rieger,M.,
            Rivolta,C., Rocha,E., Roche,B., Rose,M., Sadaie,Y., Sato,T.,
            Scanlan,E., Schleich,S., Schroeter,R., Scoffone,F., Sekiguchi,J.,
            Sekowska,A., Seror,S.J., Serror,P., Shin,B.S., Soldo,B.,
            Sorokin,A., Tacconi,E., Takagi,T., Takahashi,H., Takemaru,K.,
            Takeuchi,M., Tamakoshi,A., Tanaka,T., Terpstra,P., Tognoni,A.,
            Tosato,V., Uchiyama,S., Vandenbol,M., Vannier,F., Vassarotti,A.,
            Viari,A., Wambutt,R., Wedler,E., Wedler,H., Weitzenegger,T.,
            Winters,P., Wipat,A., Yamamoto,H., Yamane,K., Yasumoto,K., Yata,K.,
            Yoshida,K., Yoshikawa,H.F., Zumstein,E., Yoshikawa,H. and
            Danchin,A.
  TITLE     The complete genome sequence of the gram-positive bacterium
            Bacillus subtilis
  JOURNAL   Nature 390 (6657), 249-256 (1997)
  MEDLINE   98044033
   PUBMED   9384377
REFERENCE   2  (bases 1 to 197409)
  AUTHORS   Kunst,F., Ogasawara,N., Yoshikawa,H. and Danchin,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUN-2003) I. Moszer, A. Danchin, Institut Pasteur,
            Genetique des Genomes Bacteriens, 28 rue du Docteur Roux, 75724
            Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr,
            adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45
            68 89 48
COMMENT     On Jul 7, 2003 this sequence version replaced gi:2633055.
```

```
            This entry contains data from release R16.1 of the SubtiList
            database. Further data on gene annotation and detailed information
            about changes from previous releases can be found at
            http://genolist.pasteur.fr/SubtiList/.
FEATURES             Location/Qualifiers
     source          1. .197409
                     /organism="Bacillus subtilis subsp. subtilis str. 168"
                     /mol_type="genomic DNA"
                     /strain="168"
                     /db_xref="taxon:224308"
     gene            complement(52. .1608)
                     /gene="yfmM"
                     /locus_tag="BSU07420"
     CDS             complement(52. .1608)
                     /gene="yfmM"
                     /locus_tag="BSU07420"
                     /function="unknown"
                     /note="similar to ABC transporter (ATP-binding protein)"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAB12571.1"
                     /db_xref="GI:2633066"
                     /db_xref="GOA:O34512"
                     /db_xref="UniProt/TrEMBL:O34512"
                     /translation="MSILSVKDLSHGFGDRAIFNNVSFRLLKGEHVGLIGANGEGKST
                     FMNIITGKLEPDEGKVEWSKNVRVGYLDQHTVLEKGKSIRDVLKDAFHYLFAMEEEMN
                     EIYNKMGEADPDELEKLLEEVGVIQDALTNNDFYVIDSKVEEIARGLGLSDIGLERDV
                     TDLSGGQRTKVLLAKLLLEKPEILLLDEPTNYLDEQHIEWLKRYLQEYENAFILISHD
                     IPFLNSVINLIYHVENQELTRYVGDYHQFMEVYEVKKQQLEAAYKKQQQEVAELKDFV
                     ARNKARVSTRNMAMSRQKKLDKMDMIELAAEKPKPEFHFKPARTSGKLIFETKDLVIG
                     YDSPLSRPLNLRMERGQKIALYGANGIGKTTLLKSLLGEIQPLEGSVERGEHIYTGYF
                     EQEVKETNNNTCIEEVWSEFPSYTQYEIRAAPAKCGLTTKHIESRVSVLSGGEKAKVR
                     LCKLINSETNLLVLDEPTNHLDADAKEELKRALKEYKGSILLISHEPDFYMDIATETW
                     NCESWTTKVL"
     gene            1781. .2911
                     /gene="yfmL"
                     /locus_tag="BSU07430"
     CDS             1781. .2911
                     /gene="yfmL"
                     /locus_tag="BSU07430"
                     /function="unknown"
                     /note="similar to RNA helicase"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAB12572.1"
                     /db_xref="GI:2633067"
                     /db_xref="GOA:O34750"
                     /db_xref="UniProt/TrEMBL:O34750"
                     /translation="MTQTWPFLHNAQSFIQENWNASGFQKPTPVQEQAAQLIMDGKDV
                     IAESPTGTGKTLAYALPVLERIKPEQKHPQAVILAPSRELVMQIFQVIQDWKAGSELR
                     AASLIGGANVKKQVEKLKKHPHIIVGTPGRVFELIKAKKLKMHEVKTIVLDETDQLVL
                     PEHRETMKQIIKTTLRDRQLLCFSATLKKETEDVLRELAQEPEVLKVQRSKAEAGKVK
                     HQYLICDQRDKVKLLQKLSRLEGMQALVFVRDIGNLSVYAEKLAYHHVELGVLHSEAK
                     KMERAKIIATFEDGEFPLLLATDIAARGLDIENLPYVIHADIPDEDGYVHRSGRTGRA
                     GKEGNVLSLVTKLEESKLKKMAKKLGVELSEAVYAGGKLKTK"
     gene            2979. .3425
                     /gene="yfmK"
                     /locus_tag="BSU07440"
     CDS             2979. .3425
                     /gene="yfmK"
                     /locus_tag="BSU07440"
                     /function="unknown"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAB12573.1"
                     /db_xref="GI:2633068"
                     /db_xref="GOA:O34536"
                     /db_xref="UniProt/TrEMBL:O34536"
                     /translation="MASIDRFQVMQEPELKILEHWFENEDTRRRMDGMLLLDEWYARV
                     NKDKHDTVIMAYDGQLPAGMVVIEFGEERTYIGLIVNPLYRLKGYGKQILQKLMTEPD
                     FTSVREWVACIEEDNRISLACFQAAGFTLEDTEPDEDGFLTLILRN"
     gene            complement(3478. .4497)
                     /gene="yfmJ"
```

```
                     /locus_tag="BSU07450"
     CDS             complement(3478. .4497)
                     /gene="yfmJ"
                     /locus_tag="BSU07450"
                     /function="unknown"
                     /note="similar to quinone oxidoreductase"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAB12574.1"
                     /db_xref="GI:2633069"
                     /db_xref="GOA:O34812"
                     /db_xref="UniProt/TrEMBL:O34812"
                     /translation="MTASQQQIQLARRPQGIPVHEDFRFETIPVPEPKQGEVLVKTLY
                     VSVDPYMRGRMQDTKSYVEPFALDKALSGGVIAEVVSDGNHLKKGDIVIGNLSWQEFS
                     AVSESALRKIDTSLAPASAYLGILGMTGLTAYFGLLDIGRPKEGETVVVSGAAGAVGS
                     TVGQIAKIKGARVVGIAGSDEKIDYLKQELQFDEAINYKTADDIQKALQNACPDGVDV
                     YFDNVGGPISDAVMNLLNEFARIPVCGAISSYNAESEADDMGPRVQSKLIKTKSLMQG
                     FIVSDYSDRFSEGAKQLAEWLKAGKLHYEETITEGFENIPDAFLGLFKGENKGKQLIK
                     VSDPS"
     gene            complement(4979. .6199)
                     /gene="yfmI"
                     /locus_tag="BSU07460"
     CDS             complement(4979. .6199)
                     /gene="yfmI"
                     /locus_tag="BSU07460"
                     /function="unknown"
                     /note="similar to macrolide-efflux transporter"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAB12575.1"
                     /db_xref="GI:2633070"
                     /db_xref="UniProt/TrEMBL:O34440"
                     /translation="MSKGYYSILFSQTTTNLGFSLYTMVVISFLFKMTNSTTIASMVT
                     LISIIFRIFGSAILPLITIRLKLPLTIIISQLIQMLLLICLFYALLRTYSNQTLILVF
                     IIISCISFFNGWFSPLKSAIIGEIISPDRRVKANGLLSTVDQTFQFVGWSLGGLIIAF
                     LGEGYTIILTIFLLFTSLLSLLFCLPHGHANSVIREKNSPNKSIVSGWRYLFSQKKLR
                     TIIIMDLIESWAGMIWIGSVSLAFVNEVLHKGESWWGFINGAYYLGSMIGGFIIYKLS
                     ERFQNKLINFMLIGAVSYGSLTLIYGFISNSYLALILVLFMGPAYILRDLTQETLIQN
                     ITTEQTRINIMSARSSLVQFIFMFSILAIGAISDFLGVRLVYVSAGILLLVSAIYGFS
                     QLQFKKKVNKHISF"
     gene            6397. .6468
                     /gene="yfmH"
                     /locus_tag="BSU07470"
     CDS             6397. .6468
                     /gene="yfmH"
                     /locus_tag="BSU07470"
                     /function="unknown"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAB12576.1"
                     /db_xref="GI:2633071"
                     /db_xref="UniProt/TrEMBL:O34890"
                     /translation="MRKLRKNWRKSIVSPVLYTNEKW"
     gene            6535. .7998
                     /gene="yfmG"
                     /locus_tag="BSU07480"
     CDS             6535. .7998
                     /gene="yfmG"
                     /locus_tag="BSU07480"
                     /function="unknown"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAB12577.1"
                     /db_xref="GI:2633072"
                     /db_xref="GOA:O34722"
                     /db_xref="UniProt/TrEMBL:O34722"
                     /translation="MTLLMQDHKSIYNKDIEILIEVEELITSVPENKKNDFKSLIDNR
                     IMTLLNKIQHPVAAENVFGMCIAYPNTYIRYRLIDFVEKWVPYFSAVETIINLTQDPD
                     DLVSFKAMDVCANHKIEESVAYLSSIIDDVRESISYPKKPVGLGAQKVLSTLLDIFGV
                     EKHEELVMLKNYFDQNGILPNNFDFEEKIPQSLIEEFEKTEEDGMILIPGGFFEFGLN
                     ENEIPDKTFNWKDAVPRQKVWLPPFFIDKYPVTNKDYDIFTEFIEENGHIFCHPNEPQ
                     NKQHRRNTYWDDRYLDNHPVTGIDFYDAFAYARYKGKELPTEFQWEKAARGEKGNVWP
                     WGDKFDPAKVQYAGSLYNEPITSLKSWRENLLKAHADKELNHLTSDIFEQNGESPYGV
```

```
  Query Match             18.9%;  Score 190.8;  DB 1;  Length 197409;
  Best Local Similarity   55.0%;  Pred. No. 2.6e-20;
  Matches  488;  Conservative    0;  Mismatches  357;  Indels   43;  Gaps    4;

Qy          115 TCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCT 174
                || || |||      |  ||| |  |||||||||| |  |||| ||||         |||
Db        12288 TCTAAGGATTCCAGAGTGATCCATGACGAAGAAGGAAAAACGACAGTAAGCGGCACACCT 12229

Qy          175 AAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAA 234
                || || || ||||| ||||||  ||||||  | |||||  |      | | |   |||
Db        12228 AAGCGGGTGGTTGTGCTTGAGCTTTCATTCTTGGATGCCGTTCACAATCTCGGCATTACG 12169

Qy          235 CCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAA 294
                || || || || || ||||| ||||||||| ||   ||||  || |   |  |
Db        12168 CCGGTGGGCATCGCAGATGACAACAAAAAAGATATGATTAAAAAGCTTGTCGGCAGCTCC 12109

Qy          295 ATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAA 354
                ||||   | ||||| |||||||| |||||     |||| ||  | || |  ||||||
Db        12108 ATTG---ATTACACATCTGTAGGCACACGCAGCGAACCCAATCTTGAGGTCATCAGTTCC 12052

Qy          355 CTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTA 414
                 | || || |||||||| || |||||   | |  | || |||   ||||||||| |  |
Db        12051 TTGAAGCCTGATTTAATCATCGCTGACGCTGAGCGCCATAAAAACATTTATAAACAGCTG 11992

Qy          415 AATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATT 474
                || ||||| || || ||||||||| | |||||    || |     ||| | ||||  ||
Db        11991 AAAAAAATCGCCCCGACGATTGAATTAAAAAGCCGTGAAGCGACATATGACGAAACGATC 11932

Qy          475 GATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTA 534
                ||    ||||  || ||| | ||||| |||  ||||||||| ||||| ||| ||    |
Db        11931 GACAGCTTTACGACCATTGCTAAAGCATTAAATAAAGAAGATGAAGGAAAAGAAAAGCTT 11872

Qy          535 GAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAA 594
                |  || ||| | ||    ||  | ||    |||   ||| |  | |  ||| |||| |
Db        11871 GCCGAGCACAAAAAAGTCATCAACGATCTAAAAGCCGAACTTCCGAAAGATGAAAACCGC 11812

Qy          595 AAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTAT 654
                ||  |  | |  |  || ||             || |    |||| ||      || |||
Db        11811 AACATCGTTCTCGGCGTTGCAAGAGCGGATTCCTTCCAGCTTCATACATCATCATCCTAT 11752

Qy          655 GTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAA 714
                |  ||  || || | |     ||||| |||| | | ||  | |     |||
Db        11751 GACGGAGAAATCTTTAAAATGCTAGGCTTTACACACGCTGTGAAGTCAGAT--------- 11701

Qy          715 GGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAA 774
                                         | ||   | |  | | |  |||    ||     ||
Db        11700 ---------------------AACGCCTATCAAGAGGTCAGCCTTGAGCAATTGAGCAAA 11662

Qy          775 GTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
                 |  |||||||   | |||||||     | ||| |||  |   | |  |         ||
Db        11661 ATCGATCCTGATATTTTGTTCATCTCAGCCAACGAAGGCAAAACCATTG---------TA 11611

Qy          835 AAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTT 894
                 | ||     |||   | ||  | ||||| ||  | || || |||||||||  ||  ||
Db        11610 GATGAGTGGAAAACGAACCCGCTCTGGAAAAATCTCAAAGCGGTGAAAAATGGACAAGTC 11551

Qy          895 GATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATG 954
                 ||  |   |||||||||   ||| ||||||   | ||  | |  |||   |||  ||
Db        11550 TATGATGCGGACCGTGACACTTGGACAAGATTCAGAGGCATCAAGTCTAGTGAAACAAGC 11491

Qy          955 GCAAAAGAACTTGTTGAATTA-TCTAAGAAAGATAGTAAAAAAGATAA 1001
                || |||||  |  || ||  | | ||  | | |||| | ||| || ||
Db        11490 GCCAAAGATGTGCTTAAAAAAGTGTATAATAAATAGGACAAATGAGAA 11443


RESULT 21
AE017333_11
WPCOMMENT
   Sequence split into 43 fragments  LOCUS AE017333 Accession AE017333
      Fragment Name       Begin      End
      AE017333_00             1    110000
      AE017333_01        100001    210000
```

```
      AE017333_02        200001    310000
      AE017333_03        300001    410000
      AE017333_04        400001    510000
      AE017333_05        500001    610000
      AE017333_06        600001    710000
      AE017333_07        700001    810000
      AE017333_08        800001    910000
      AE017333_09        900001   1010000
      AE017333_10       1000001   1110000
      AE017333_11       1100001   1210000
      AE017333_12       1200001   1310000
      AE017333_13       1300001   1410000
      AE017333_14       1400001   1510000
      AE017333_15       1500001   1610000
      AE017333_16       1600001   1710000
      AE017333_17       1700001   1810000
      AE017333_18       1800001   1910000
      AE017333_19       1900001   2010000
      AE017333_20       2000001   2110000
      AE017333_21       2100001   2210000
      AE017333_22       2200001   2310000
      AE017333_23       2300001   2410000
      AE017333_24       2400001   2510000
      AE017333_25       2500001   2610000
      AE017333_26       2600001   2710000
      AE017333_27       2700001   2810000
      AE017333_28       2800001   2910000
      AE017333_29       2900001   3010000
      AE017333_30       3000001   3110000
      AE017333_31       3100001   3210000
      AE017333_32       3200001   3310000
      AE017333_33       3300001   3410000
      AE017333_34       3400001   3510000
      AE017333_35       3500001   3610000
      AE017333_36       3600001   3710000
      AE017333_37       3700001   3810000
      AE017333_38       3800001   3910000
      AE017333_39       3900001   4010000
      AE017333_40       4000001   4110000
      AE017333_41       4100001   4210000
      AE017333_42       4200001   4222645
Continuation (12 of 43) of AE017333 from base 1100001 (AE017333 Bacillus licheniformis

  Query Match             16.7%;  Score 168.4;  DB 1;  Length 110000;
  Best Local Similarity   50.5%;  Pred. No. 7.7e-17;
  Matches  490;  Conservative    0;  Mismatches  471;  Indels    9;  Gaps    3;

Qy           41 TTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACT 100
                |||   | || ||| | || ||   | | |        |       | |   |||  |
Db        11539 TTGCTATTTTATTTTTATTGATGATATCAGTGATGACGGCCTGCTCTGGAGGAAGCGAGA 11598

Qy          101 CAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAG 160
                  |  |||   ||     | |||   ||   | | ||||| ||    ||    |||||||
Db        11599 GCACGAAAACTTCCGGTCAGGATTCCGTGACAGTTAAGCATGATCTTGGAGAAACGAAAG 11658

Qy          161 TACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTG 220
                | |  ||  | || |||   ||||| |||||||| |    |||| |||||||| |  |||
Db        11659 TGCGGAATAAACCGAAAAAAGTTGTGGTTCTTGAATTGGGATTTATTGATGCGCTCCTTG 11718

Qy          221 CTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAAC 280
                 |   | | |||| ||||| || || ||||| ||   ||||   ||   |||||   |
Db        11719 ATGCGGGTATTAAGCCTGTCGGAATTGCGGACGACGGCAAACCTAAGTTTATTAACGAGA 11778

Qy          281 CATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAG 340
                   | || |  |||||   |   || ||||| || ||  | ||    || || | ||| |
Db        11779 AGGTCAGGGGAAAAATCAAAGGGTATACTTCAGTCGGTTCGCGCGCCCAGCCAAGCTTTG 11838

Qy          341 AAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTA 400
                || ||||   |    | || || |||||||||||||| |||   |  || ||||  ||
Db        11839 AAAAAATTGCTTCTTTAAAGCCCGATTTAATTATTGCCGATTCGAGCAGGCACAGCGGCG 11898

Qy          401 TTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATT 460
                | ||| |  |  |    |||||||| || || || |  || ||||  ||  |||  ||||
Db     11899 TCTATGATAAGCTGTCGAAAATTGCGCCGACAATCGCGCTCAAAAACTTGAATGCCGATT 11958

Qy       461 ATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAG 520
             || | || |   ||||||| | |   || ||| | |||||  | || ||||||   | |
Db     11959 ATCAGGACACGCTTGATGCATCTCTTACGATTGCGAAAGCCGTCGGCAAAGAAAGCGCAA 12018

Qy       521 GTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTA 580
                | |||   || |  |||||  |  | ||  | || ||||  ||| |  || |    |
Db     12019 TGGAGAAAAAATTGGCTGAACATAAACAAAAGCTGGATGAATTGAAACAGAAATTCGGCA 12078

Qy       581 TGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATC 640
                  ||| |    |      |||  |      |   | ||  |       |     ||
Db     12079 GCCGGAAACAGAGCATTCTTCTGCTCGGGAACACAAATGAAGAAATCACCGTGCGCGATG 12138

Qy       641 CAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTG 700
              |   |||| ||       |||  | || |   || | || | || | | |   |  | |
Db     12139 AA---AACTTTTTCACGTCTCAGCTTCTGACGAAAATCGGCTATACATACGGTGTCGGAG 12195

Qy       701 ATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTG 760
             |    |     |||||        |  |   | ||       |  |  ||||||     |
Db     12196 ACAGCG---GCAAAGGCGATGCCGAAAACGGTGAATCCGTCAATATTAAAATGACGCTCG 12252

Qy       761 AAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGAC---AAACAAAGCAAGTT 817
             ||    |   | |     |||| ||   |||  || | |||||   | | ||||  |
Db     12253 AACAGCTGCTTGAGAAAGATCCCGACGTTATTGTCCTGATGACGGGAGAAAAAGATAAAG 12312

Qy       818 CTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTG 877
                | ||| |      || |    | |||||||||||| | ||||| ||  | | ||| |
Db     12313 TCGATGAAGACGGGAAAAGACCGATCGAAAAAGATCCTCTTTGGAAAAAGCTCAGCGCAG 12372

Qy       878 TGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAA 937
             | |||||       ||| ||      ||| |   |   ||| |      |||  |  |
Db     12373 TCAAAAACGGCAAGGTTTATGAAGCCGACAGATTCGCCTGGTCGCTCCGACGCAGCATTG 12432

Qy       938 TTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAG 997
                   |||| ||  ||    | ||| | | | |      |  |   ||   | ||||
Db     12433 ACGGAGCAGATGAGCTGATGGATGAAATCGATCAGCATCTTCTGTCTGAGGATCAAAAGA 12492

Qy       998 ATAATAAGTA 1007
             |   | || |
Db     12493 AATCTTAGGA 12502


RESULT 22
CP000002_11
WPCOMMENT
  Sequence split into 43 fragments  LOCUS CP000002 Accession CP000002
     Fragment Name        Begin      End
     CP000002_00              1    110000
     CP000002_01         100001    210000
     CP000002_02         200001    310000
     CP000002_03         300001    410000
     CP000002_04         400001    510000
     CP000002_05         500001    610000
     CP000002_06         600001    710000
     CP000002_07         700001    810000
     CP000002_08         800001    910000
     CP000002_09         900001   1010000
     CP000002_10        1000001   1110000
     CP000002_11        1100001   1210000
     CP000002_12        1200001   1310000
     CP000002_13        1300001   1410000
     CP000002_14        1400001   1510000
     CP000002_15        1500001   1610000
     CP000002_16        1600001   1710000
     CP000002_17        1700001   1810000
     CP000002_18        1800001   1910000
     CP000002_19        1900001   2010000
     CP000002_20        2000001   2110000
     CP000002_21        2100001   2210000
     CP000002_22        2200001   2310000
     CP000002_23        2300001   2410000
     CP000002_24        2400001   2510000
     CP000002_25        2500001   2610000
     CP000002_26        2600001   2710000
     CP000002_27        2700001   2810000
     CP000002_28        2800001   2910000
     CP000002_29        2900001   3010000
     CP000002_30        3000001   3110000
     CP000002_31        3100001   3210000
     CP000002_32        3200001   3310000
     CP000002_33        3300001   3410000
     CP000002_34        3400001   3510000
     CP000002_35        3500001   3610000
     CP000002_36        3600001   3710000
     CP000002_37        3700001   3810000
     CP000002_38        3800001   3910000
     CP000002_39        3900001   4010000
     CP000002_40        4000001   4110000
     CP000002_41        4100001   4210000
     CP000002_42        4200001   4222336
Continuation (12 of 43) of CP000002 from base 1100001 (CP000002 Bacillus licheniformis D

  Query Match             16.7%;  Score 168.4;  DB 1;  Length 110000;
  Best Local Similarity   50.5%;  Pred. No. 7.7e-17;
  Matches  490;  Conservative    0;  Mismatches  471;  Indels    9;  Gaps    3;

Qy        41 TTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACT 100
             |||   | || ||| | || ||   | | |        |       | |   |||  |
Db     10685 TTGCTATTTTATTTTTATTGATGATATCAGTGATGACGGCCTGCTCTGGAGGAAGCGAGA 10744

Qy       101 CAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAG 160
               |  |||   ||     | |||   ||   | | ||||| ||    ||    |||||||
Db     10745 GCACGAAAACTTCCGGTCAGGATTCCGTGACAGTTAAGCATGATCTTGGAGAAACGAAAG 10804

Qy       161 TACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTG 220
             | |  ||  | || |||   ||||| |||||||| |    |||| |||||||| |  |||
Db     10805 TGCGGAATAAACCGAAAAAAGTTGTGGTTCTTGAATTGGGATTTATTGATGCGCTCCTTG 10864

Qy       221 CTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAAC 280
              |   | | |||| ||||| || || ||||| ||   ||||   ||   |||||   |
Db     10865 ATGCGGGTATTAAGCCTGTCGGAATTGCGGACGACGGCAAACCTAAGTTTATTAACGAGA 10924

Qy       281 CATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAG 340
                | || |  |||||   |   || ||||| || ||  | ||    || || | ||| |
Db     10925 AGGTCAGGGGAAAAATCAAAGGGTATACTTCAGTCGGTTCGCGCGCCCAGCCAAGCTTTG 10984

Qy       341 AAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTA 400
             || ||||   |    | || || |||||||||||||| |||   |  || ||||  ||
Db     10985 AAAAAATTGCTTCTTTAAAGCCCGATTTAATTATTGCCGATTCGAGCAGGCACAGCGGCG 11044

Qy       401 TTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATT 460
             | ||| |  |  |    |||||||| || || || |  || ||||  ||  |||  ||||
Db     11045 TCTATGATAAGCTGTCGAAAATTGCGCCGACAATCGCGCTCAAAAACTTGAATGCCGATT 11104

Qy       461 ATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAG 520
             || | || |   ||||||| | |   || ||| | |||||  | || ||||||   | |
Db     11105 ATCAGGACACGCTTGATGCATCTCTTACGATTGCGAAAGCCGTCGGCAAAGAAAGCGCAA 11164

Qy       521 GTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTA 580
                | |||   || |  |||||  |  | ||  | || ||||  ||| |  || |    |
Db     11165 TGGAGAAAAAATTGGCTGAACATAAACAAAAGCTGGATGAATTGAAACAGAAATTCGGCA 11224

Qy       581 TGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATC 640
                  ||| |    |      |||  |      |   | ||  |       |     ||
Db     11225 GCCGGAAACAGAGCATTCTTCTGCTCGGGAACACAAATGAAGAAATCACCGTGCGCGATG 11284

Qy       641 CAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTG 700
              |   |||| ||       |||  | || |   || | || | || | | |   |  | |
Db     11285 AA---AACTTTTTCACGTCTCAGCTTCTGACGAAAATCGGCTATACATACGGTGTCGGAG 11341

Qy       701 ATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTG 760
             |    |     |||||        |  |   | ||       |  |  ||||||     |
Db     11342 ACAGCG---GCAAAGGCGATGCCGAAAACGGTGAATCCGTCAATATTAAAATGACGCTCG 11398

Qy       761 AAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGAC---AAACAAAGCAAGTT 817
```

```
              ||    |   | |     |||| ||   |||  || | |||||   | | ||||  |
Db      11399 AACAGCTGCTTGAGAAAGATCCCGACGTTATTGTCCTGATGACGGGAGAAAAAGATAAAG 11458

Qy        818 CTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTG 877
                 | |||        || |    | |||||||||||| | ||||| ||  | | ||| |
Db      11459 TCGATGAAGACGGGAAAAGACCGATCGAAAAAGATCCTCTTTGGAAAAAGCTCAGCGCAG 11518

Qy        878 TGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAA 937
              | |||||       ||| ||      ||| |   |   ||| |      |||  |  |
Db      11519 TCAAAAACGGCAAGGTTTATGAAGCCGACAGATTCGCCTGGTCGCTCCGACGCAGCATTG 11578

Qy        938 TTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAG 997
                    |||| ||  ||    | ||| | | | |      |  |   ||   | ||||
Db      11579 ACGGAGCAGATGAGCTGATGGATGAAATCGATCAGCATCTTCTGTCTGAGGATCAAAAGA 11638

Qy        998 ATAATAAGTA 1007
              |   | || |
Db      11639 AATCTTAGGA 11648


RESULT 23
BSY14084
LOCUS       BSY14084               12556 bp    DNA     linear   BCT 13-OCT-2003
DEFINITION  Bacillus subtilis chromosomal DNA, region 78-80 degrees: aprE to
            comK.
ACCESSION   Y14084
VERSION     Y14084.1  GI:2226250
KEYWORDS    comK gene; yhfO gene; yhfP gene; yhfQ gene; yhfR gene; yhfS gene;
            yhfT gene; yhfU gene; yhfV gene; yhfW gene; yhxC gene; yhxD gene.
SOURCE      Bacillus subtilis
  ORGANISM  Bacillus subtilis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE   1  (bases 1 to 104)
  AUTHORS   Stahl,M.L. and Ferrari,E.
  TITLE     Replacement of the Bacillus subtilis subtilisin structural gene
            with an In vitro-derived deletion mutation
  JOURNAL   J. Bacteriol. 158 (2), 411-418 (1984)
  MEDLINE   84212198
   PUBMED   6427178
REFERENCE   2  (bases 1 to 441)
  AUTHORS   Park,S.S., Wong,S.L., Wang,L.F. and Doi,R.H.
  TITLE     Bacillus subtilis subtilisin gene (aprE) is expressed from a sigma
            A (sigma 43) promoter in vitro and in vivo
  JOURNAL   J. Bacteriol. 171 (5), 2657-2665 (1989)
  MEDLINE   89213955
   PUBMED   2496113
REFERENCE   3  (bases 10610 to 12556)
  AUTHORS   van Sinderen,D., ten Berge,A., Hayema,B.J., Hamoen,L. and Venema,G.
  TITLE     Molecular cloning and sequence of comK, a gene required for genetic
            competence in Bacillus subtilis
  JOURNAL   Mol. Microbiol. 11 (4), 695-703 (1994)
  MEDLINE   94254725
   PUBMED   8196543
REFERENCE   4  (bases 1 to 12556)
  AUTHORS   Noback,M.A., Terpstra,P., Holsappel,S., Venema,G. and Bron,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUN-1997) Noback M. A., University of Groningen,
            Dept. of Genetics, Kerklaan 30, Haren, 9751 NN The Netherlands
FEATURES             Location/Qualifiers
     source          1. .12556
                     /organism="Bacillus subtilis"
                     /mol_type="genomic DNA"
                     /strain="168"
                     /db_xref="taxon:1423"
                     /map="78 to 80 degrees, aprE to comK"
     gene            383. .852
                     /gene="yhfO"
     RBS             383. .396
                     /gene="yhfO"
     CDS             403. .852
                     /gene="yhfO"
                     /note="Similarity to several hypothetical proteins, from
                     Haemophilus influenzae (Swiss Prot P44036), Saccharomyces
```

```
                     cerevisiae (Swiss Prot P39979), a
                     /citation=[2]
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="CAA74538.1"
                     /db_xref="GI:2226251"
                     /db_xref="GOA:O07614"
                     /db_xref="UniProt/TrEMBL:O07614"
                     /translation="MNIRQAKTSDAAAIAPLFNQYREFYRQA
                     HESVILIAEENGEFIGFTQLYPTFSSVSMKRIYILNDLFVVP              AAK
                     DYAGQNGAKCLTLQTEHHNRKARSLYEQNGYEEDTGFVHYCLN
     gene            907. .1916
                     /gene="yhfP"
     RBS             907. .919
                     /gene="yhfP"
     CDS             924. .1916
                     /gene="yhfP"
                     /note="high similarity to yhdH from E.coli (Swiss Prot
                     P26646), and lower to many alcohol dehydrogenases."
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="CAA74539.1"
                     /db_xref="GI:2226252"
                     /db_xref="GOA:O07615"
                     /db_xref="UniProt/TrEMBL:O07615"
                     /translation="MSTLFQALQAEKNADDVSVHVKTISTEDLPKDGVLIKVAYSGIN
                     YKDGLAGKAGGNIVREYPLILGIDAAGTVVSSNDPRFAEGDEVIATSYELGVSRDGGL
                     SEYASVPGDWLVPLPQNLSLKEAMVYGTAGFTAALSVHRLEQNGLSPEKGSVLVTGAT
                     GGVGGIAVSMLNKRGYDVVASTGNREAADYLKQLGASEVISREDVYDGTLKALSKQQW
                     QGAVDPVGGKQLASLLSKIQYGGSVAVSGLTGGGEVPATVYPFILRGVSLLGIDSVYC
                     PMDVRAAVWERMSSDLKPDQLLTIVDREVSLEETPGALKDILQNRIQGRVIVKL"
     gene            2043. .3104
                     /gene="yhfQ"
     RBS             2043. .2056
                     /gene="yhfQ"
     CDS             2058. .3104
                     /gene="yhfQ"
                     /note="Similarity to citrate-dependent iron transport
                     protein fecB from E.coli (PIR S56515) and Synechocystis
                     sp. (EMBL: D90899)."
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="CAA74540.1"
                     /db_xref="GI:2226253"
                     /db_xref="GOA:O07616"
                     /db_xref="UniProt/TrEMBL:O07616"
                     /translation="MILIPCYTTLGHLNHRNYQQERLEHMKKTLIILTVLLLSVLTAA
                     CSSSSGNQNSKEHKVAVTHDLGKTNVPEHPKRVVVLELGFIDTLLDLGITPVGVADDN
                     KAKQLINKDVLKKIDGYTSVGTRSQPSMEKIASLKPDLIIADTTRHKKVYDQLKKIAP
                     TIALNNLNADYQDTIDASLTIAKAVGKEKEMEKKLTAHEEKLSETKQKISANSQSVLL
                     IGNTNDTIMARDENFFTSRLLTQVGYRYAISTSGNSDSSNGGDSVNMKMTLEQLLKTD
                     PDVIILMTGKTDDLDADGKRPIEKNVLWKKLKAVKNGHVYHVDRAVWSLRRSVDGANA
                     ILDELQKEMPAAKK"
     terminator      3099. .3141
     gene            complement(3136. .3731)
                     /gene="yhfR"
     CDS             complement(3136. .3717)
                     /gene="yhfR"
                     /note="similarity to a hypothetical protein o215b from
                     E.coli (Swiss Prot P36942), and to several
                     phosphoglycerate mutase (pgm) proteins, from Treponema
                     pallidum (EMBL: U55214) and Schizosaccharomyces pombe
                     (Swiss Prot P36623)."
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="CAA74541.1"
                     /db_xref="GI:2226254"
                     /db_xref="GOA:O07617"
                     /db_xref="UniProt/TrEMBL:O07617"
```

1): BCT 07-JUL-2003
from 1011039

```
                     /translation="MTAVCLVRHGETDWNLQQKCQGKTDIPLNATGERQARETGEYVK
                     DFSWDIIVTSPLKRAKRTAEIINEYLHLPIVEMDDFKERDYGDAEGMPLEERTKRYPD
                     NIYPNMETLEELTDRLMGGLAKVNQAYPNKKVLIVAHGAAIHALLTEISGGDPELQST
                     RLVNACLSNIEFAEEKWRIKDYNINSHLSGFIK"
     RBS             complement(3722. .3731)
                     /gene="yhfR"
     gene            complement(3788. .4882)
                     /gene="yhfS"
     CDS             complement(3788. .4882)
                     /gene="yhfS"
                     /note="similarity to acetyl-CoAuacetyltransferases (ThiL,
                     encoding the enzyme for the first step in
                     poly-beta-hydroxybutyrate synthesis) from Chromatium
                     vinosum (Swiss Prot P45369), Alcaligenes eutrophus (Swiss
                     Prot P14611), Thiocystis violacea (Swiss Prot P45363), and
                     B.subtilis (Swiss Prot P45855)."
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="CAA74542.1"
                     /db_xref="GI:2226255"
                     /db_xref="UniProt/TrEMBL:O07618"
                     /translation="MNAVIVDAKRTIFGNQNGLLKPFLPEDLAAPIIRCLSRKLEDQV
                     DEVILGNATGRGGNLARLSALQAGLPLSVPGMTIDRQCGSGLEAVRYACSLIQAGAGT
                     MYIAGGSESSSQSPFSERARFSPDAIGDPDMGIAAEYTAARYSISRSMQDEYALLSHQ
                     RSRNAHDEGFYREEVVALGELETDEAFLKTRPIEAIIPRAKPVFDTSSGTVTAANSSG
                     IADGAAALLVMEEEKAAALGLKPVLRFIGSAVSGIHPNFPPAAPVVAIRQLLHTHDVT
                     PDDIDLFEINEAFAVKICVCSQELGIPFSKINVRGGALALGHPYGASGAALVTRLFYE
                     AKRRPDCQYAVAAIGSGGGIGLALLFEVLA"
     gene            complement(4879. .6318)
                     /gene="yhfT"
     CDS             complement(4879. .6318)
                     /gene="yhfT"
                     /note="similarity to long-chain-acyl-CoA synthetase from
                     B.subtilis (EMBL: Z75208), bile acid-CoA ligase from
                     Eubacterium sp. (Swiss Prot P19409), long-chain-fatty
                     aciduCoA ligase from E.coli (LcfA; Swiss Prot P29212)."
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="CAA74543.1"
                     /db_xref="GI:2226256"
                     /db_xref="GOA:O07619"
                     /db_xref="UniProt/TrEMBL:O07619"
                     /translation="MTITHTYSSTAETSPGRVAIQTESEQITYHDWDRLVSQTANWLR
                     SQPSMPNRVAILLPNSLAFLQLFAGAAAAGCTAIPIDTRWSPAECKERLSISNADLVV
                     TLAFFKNKLTDSQTPVVLLDNCMADISEAAADPLPTIDPEHPFYMGFTSGSTGKPKAF
                     TRSHRSWMESFTCTETDFSISSDDKVLIPGALMSSHFLYGAVSTLFLGGTVCLLKKFS
                     PAKAKEWLCRESISVLYTVPTMTDALARIEGFPDSPVKIISSGADWPAESKKKLAAAW
                     PHLKLYDFYGTSELSFVTFSSPEDSKRKPHSAGRPFHNVRIEIRNAGGERCQPGEIGK
                     IFVKSPMRFSGYVNGSTPDEWMTVDDMGYVDEEGFLYISGRENGMIVYGGLNIFPEEI
                     ERVLLACPEVESAAVVGIPDEYWGEIAVAVILGNANARTLKAWCKQKLASYKIPKKWV
                     FADSLPETSSGKIARSRVKKWLEESVQYK"
     gene            complement(4893. .4903)
                     /gene="yhfS"
     RBS             complement(4893. .4903)
                     /gene="yhfS"
     gene            complement(join(6325. .6885,6888. .6897))
                     /gene="yhfU"
     CDS             complement(6325. .6885)
                     /gene="yhfU"
                     /note="similarity to BioY from Bacillus sphaericus (Swiss
                     Prot P22819)."
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="CAA74544.1"
                     /db_xref="GI:2226257"
                     /db_xref="UniProt/TrEMBL:O07620"
                     /translation="MLKLIDMMHIAIFTALMAVLGFMPPLFLSFTPVPITLQTLGVML
                     AGSILRPKSAFLSQLVFLLLVAFGAPLLPGGRGGFGVFFGPSAGFLIAYPLASWLISL
                     AANRLRKVTVLRLFFTHIVFGIIFIYLLGIPVQAFIMHIDLSQAAFMSLAYVPGDLIK
                     AAVSAFLAIKITQALSLSDTMFTKGG"
```

```
     gene            complement(6329. .6336)
                     /gene="yhfT"
     RBS             complement(6329. .6336)

  Query Match             16.5%;  Score 166.6;  DB 1;  Length 12556;
  Best Local Similarity   51.5%;  Pred. No. 2.5e-16;
  Matches  460;  Conservative    0;  Mismatches  424;  Indels    9;  Gaps    3;

Qy         117 AAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAA 176
               | || ||  ||| |   | |  || ||   ||| |  || || || ||| | || || ||
Db        2222 AGAACATAAAGTGGCGGTAACACATGATTTAGGGAAGACAAATGTGCCTGAGCATCCGAA 2281

Qy         177 ACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACC 236
                || |||||||||||||||      ||| ||||| |  |  ||| | | |   |||  ||
Db        2282 GCGGGTTGTTGTTCTTGAGCTAGGTTTTATTGATACACTGCTTGATCTCGGCATTACGCC 2341

Qy         237 TGTTGGGATAGCGGATGATAACAAA---AAAAATCGTATTATTAAACCATTAAGAGATAA 293
               ||| ||| | || ||||| ||||||   ||  | |  || |  ||     |     | ||
Db        2342 TGTCGGGGTTGCCGATGACAACAAAGCGAAGCAGCTGATCAACAAGGATGTGCTGAAGAA 2401

Qy         294 AATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAA 353
               |||||     ||||| ||||| || || ||    || || | | | ||| |||||  |
Db        2402 AATTGACGGCTACACATCTGTCGGCACTCGCTCACAGCCAAGCATGGAAAAAATCGCTTC 2461

Qy         354 ACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTT 413
               | | ||||| ||||||||||||||||| |  |   | || ||     | ||  |  |  |
Db        2462 ATTAAAACCCGATTTAATTATTGCTGACACGACCCGGCATAAGAAGGTGTACGATCAGCT 2521

Qy         414 AAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATAT 473
                || ||||| || || ||||||| ||| || | |||  |||  |||||| | || |  ||
Db        2522 GAAAAAAATAGCGCCGACGATTGCACTTAATAATTTAAATGCTGATTATCAGGATACAAT 2581

Qy         474 TGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTT 533
               ||| ||||     || ||| |||||||  | || || ||  | |||    | |||    |
Db        2582 TGACGCTTCGCTTACGATTGCAAAAGCAGTCGGCAAGGAGAAGGAAATGGAGAAAAAGCT 2641

Qy         534 AGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCA 593
                   |  || ||  | ||  ||   ||    ||  |  |||| |    | | |   |
Db        2642 GACGGCGCATGAAGAAAAGCTTAGCGAGACAAAGCAGAAAATCAGCGCGAACAGCCAGTC 2701

Qy         594 AAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTA 653
                  | | |||   | |    |   | |  |   | ||       ||  || |  || |
Db        2702 CGTGCTTTTGATCGGAAATACAAATGATACCATTATGGCCAGGGATGAAAACTTCTTTAC 2761

Qy         654 TGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAA 713
                    | |  |   |||  ||  | || |    | | ||| | |||      |  | ||
Db        2762 ATCGAGACTTT---TAACACAGGTCGGCTACCGATATGCAATCAGTACGTCAGGCAATAG 2818

Qy         714 AGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCA 773
                | || ||| ||        |        |  |  ||||||     ||     |     |
Db        2819 CGATTCAAGCAATGGCGGCGACTCTGTGAATATGAAAATGACACTGGAGCAGCTGCTGAA 2878

Qy         774 AGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACT 833
               |    |||| ||    ||  || | ||||||   ||| ||  |     |||  |
Db        2879 AACAGATCCGGATGTGATCATCCTGATGACAGGAAAAACAGATGACCTCGACGCCGACGG 2938

Qy         834 AAAAGAAC---TAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACG 890
                |||   |   | |||||  ||    | ||||| ||| | || || ||||||||    |
Db        2939 TAAACGCCCGATCGAAAAGAATGTCCTTTGGAAAAAACTGAAGGCAGTGAAAAACGGGCA 2998

Qy         891 TGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGA 950
               |||  |     | || ||||   | ||| |       ||  || ||       |  | |
Db        2999 TGTATACCACGTGGATCGTGCGGTGTGGTCTCTGCGCCGCAGTGTAGACGGGGCGAATGC 3058

Qy         951 AATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATA 1003
                ||     | || |||    ||        |   | ||  ||| || | || |
Db        3059 CATTTTGGACGAGCTTCAAAAAGAGATGCCGGCTGCTAAGAAATAAAAGAAAA 3111


RESULT 24
BSUB0006
LOCUS       BSUB0006               198743 bp    DNA     linear
DEFINITION  Bacillus subtilis complete genome (section 6 of
            to 1209781.
ACCESSION   Z99109 AL009126
VERSION     Z99109.2  GI:32468723
KEYWORDS    .
SOURCE      Bacillus subtilis subsp. subtilis str. 168
  ORGANISM  Bacillus subtilis subsp. subtilis str. 168
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE   1  (bases 1 to 198743)
  AUTHORS   Kunst,F., Ogasawara,N., Moszer,I., Albertini,A.M., Alloni,G.,
            Azevedo,V., Bertero,M.G., Bessieres,P., Bolotin,A., Borchert,S.,
            Borriss,R., Boursier,L., Brans,A., Braun,M., Brignell,S.C.,
            Bron,S., Brouillet,S., Bruschi,C.V., Caldwell,B., Capuano,V.,
            Carter,N.M., Choi,S.K., Codani,J.J., Connerton,I.F., Cummings,N.J.,
            Daniel,R.A., Denizot,F., Devine,K.M., Dusterhoft,A., Ehrlich,S.D.,
            Emmerson,P.T., Entian,K.D., Errington,J., Fabret,C., Ferrari,E.,
            Foulger,D., Fritz,C., Fujita,M., Fujita,Y., Fuma,S., Galizzi,A.,
            Galleron,N., Ghim,S.Y., Glaser,P., Goffeau,A., Golightly,E.J.,
            Grandi,G., Guiseppi,G., Guy,B.J., Haga,K., Haiech,J., Harwood,C.R.,
            Henaut,A., Hilbert,H., Holsappel,S., Hosono,S., Hullo,M.F.,
            Itaya,M., Jones,L., Joris,B., Karamata,D., Kasahara,Y.,
            Klaerr-Blanchard,M., Klein,C., Kobayashi,Y., Koetter,P.,
            Koningstein,G., Krogh,S., Kumano,M., Kurita,K., Lapidus,A.,
            Lardinois,S., Lauber,J., Lazarevic,V., Lee,S.M., Levine,A., Liu,H.,
            Masuda,S., Mauel,C., Medigue,C., Medina,N., Mellado,R.P.,
            Mizuno,M., Moestl,D., Nakai,S., Noback,M., Noone,D., O'Reilly,M.,
            Ogawa,K., Ogiwara,A., Oudega,B., Park,S.H., Parro,V., Pohl,T.M.,
            Portetelle,D., Porwollik,S., Prescott,A.M., Presecan,E., Pujic,P.,
            Purnelle,B., Rapoport,G., Rey,M., Reynolds,S., Rieger,M.,
            Rivolta,C., Rocha,E., Roche,B., Rose,M., Sadaie,Y., Sato,T.,
            Scanlan,E., Schleich,S., Schroeter,R., Scoffone,F., Sekiguchi,J.,
            Sekowska,A., Seror,S.J., Serror,P., Shin,B.S., Soldo,B.,
            Sorokin,A., Tacconi,E., Takagi,T., Takahashi,H., Takemaru,K.,
            Takeuchi,M., Tamakoshi,A., Tanaka,T., Terpstra,P., Tognoni,A.,
            Tosato,V., Uchiyama,S., Vandenbol,M., Vannier,F., Vassarotti,A.,
            Viari,A., Wambutt,R., Wedler,E., Wedler,H., Weitzenegger,T.,
            Winters,P., Wipat,A., Yamamoto,H., Yamane,K., Yasumoto,K., Yata,K.,
            Yoshida,K., Yoshikawa,H.F., Zumstein,E., Yoshikawa,H. and
            Danchin,A.
  TITLE     The complete genome sequence of the gram-positive bacterium
            Bacillus subtilis
  JOURNAL   Nature 390 (6657), 249-256 (1997)
  MEDLINE   98044033
   PUBMED   9384377
REFERENCE   2  (bases 1 to 198743)
  AUTHORS   Kunst,F., Ogasawara,N., Yoshikawa,H. and Danchin,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUN-2003) I. Moszer, A. Danchin, Institut Pasteur,
            Genetique des Genomes Bacteriens, 28 rue du Docteur Roux, 75724
            Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr,
            adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45
            68 89 48
COMMENT     On Jul 7, 2003 this sequence version replaced gi:2633260.
            This entry contains data from release R16.1 of the SubtiList
            database. Further data on gene annotation and detailed information
            about changes from previous releases can be found at
            http://genolist.pasteur.fr/SubtiList/.
FEATURES             Location/Qualifiers
     source          1. .198743
                     /organism="Bacillus subtilis subsp. subtilis str. 168"
                     /mol_type="genomic DNA"
                     /strain="168"
                     /db_xref="taxon:224308"
     gene            complement(38. .1504)
                     /gene="lytF"
                     /locus_tag="BSU09370"
     CDS             complement(38. .1504)
                     /gene="lytF"
                     /locus_tag="BSU09370"
                     /function="cell wall lytic activity"
                     /note="alternate gene name: cwlE, yhdD"
                     /codon_start=1
                     /transl_table=11
                     /product="gamma-D-glutamate-meso-diaminopimelate
```

```
                     muropeptidase (major autolysin) (CWBP49')"
                     /protein_id="CAB12776.1"
                     /db_xref="GI:2633272"
                     /db_xref="GOA:O07532"
                     /db_xref="UniProt/TrEMBL:O07532"
                     /translation="MKKKLAAGLTASAIVGTTLVVTPAEAATIKVKSGDSLWKLAQTY
                     NTSVAALTSANHLSTTVLSIGQTLTIPGSKSSTSSSTSSSTTMKSGSSVYTVKSGDSL
                     WLIANEFKMTVQELKKLNGLSSDLIRAGQKLKVSGTVSSSSSSSKKSNSNKSSSSSSK
                     SSSNKSSSSSSSTGTYKVQLGDSLWKIANKVNMSIAELKVLNNLKSDTIYVNQVLKTK
                     SSGSDTSSKDNSSKSNQTSATTKYTVKSGDSLWKIANNYNLTVQQIRNINNLKSDVLY
                     VGQVLKLTGKASSGSSSSSSSSSNASSGTTTTYTVKSGDSLWVIAQKFNVTAQQIREK
                     NNLKTDVLQVGQKLVISGKASSSSSSGSSNTTSSTSAKINTMISAAKAQLGVPYRWGG
                     TTPSGFDCSGFIYYVLNKVTSVSRLTAAGYWNTMKSVSQPAVGDFVFFSTYKAGPSHV
                     GIYLGNGEFINANDSGVVISNMNNSYWKQRYLGAKRYF"
     gene            complement(1657. .2097)
                     /gene="yhdE"
                     /locus_tag="BSU09380"
     CDS             complement(1657. .2097)
                     /gene="yhdE"
                     /locus_tag="BSU09380"
                     /function="unknown"
                     /note="similar to unknown proteins"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAB12777.1"
                     /db_xref="GI:2633273"
                     /db_xref="GOA:O07573"
                     /db_xref="UniProt/Swiss-Prot:O07573"
                     /translation="MKLTNYTDYSLRVLIFLAAERPGELSNIKQIAETYSISKNHLMK
                     VIYRLGQLGYVETIRGRGGGIRLGMDPEDINIGEVVRKTEDDFNIVECFDANKNLCVI
                     SPVCGLKHVLNEALLAYLAVLDKYTLRDLVKNKEDIMKLLKMKE"
     gene            complement(2200. .3858)
                     /gene="ygxB"
                     /locus_tag="BSU09390"
     CDS             complement(2200. .3858)
                     /gene="ygxB"
                     /locus_tag="BSU09390"
                     /function="unknown"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAB12778.1"
                     /db_xref="GI:2633274"
                     /db_xref="GOA:P37874"
                     /db_xref="UniProt/Swiss-Prot:P37874"
                     /translation="MYEPARQKMIVSLGQEKIPPAHSSVCLLDKWVNTHHTTKKERYQ
                     LNVADVFSSYLSKLPNVIIALLVLLIGWAIAKIIEKAVYKGLSKTKIDDKLFAGKKPS
                     RYSSEKVISKVVYFIALIIVFILFFNILHLTTVASPFVSMLSAIAAAIPSVLKAGLIL
                     LLGWAAASVLSFLVKKIGMKLNTSDKLRKWNLVSEGKDIHQAVNTASQIVFYLVLLVF
                     LPGVLSSLKISGISGPFTNMMESVLAFLPKLFAAALIVLIGWLVARLVRDIITNFLAS
                     IGTERFAARMGLSIYLKDTSLSAVIGTIAYVLIMIPVVISALDQLDVAGISKPAVSML
                     NTILNMLPNIMIAIVLVLAGIWAGKWVKSMVSGLLHRAGFDSVLGKMGMEAGTPAKLS
                     LSQVVGMIAQIIVILLFTAEALQIVRLHFLVEIATGIIAYLPNVLVAVFILGLGLYAG
                     ELVRKVLASMIKGQEFKSLAPIAKYTIIALAFFMALDQLGVAATIVNSAFIIVLSGFA
                     LAFGLSFGLGGKDFASRYLSTFERKMQNTEIEKNRKNQNPPNDM"
     gene            3889. .5295
                     /gene="spoVR"
                     /locus_tag="BSU09400"
     CDS             3889. .5295
                     /gene="spoVR"
                     /locus_tag="BSU09400"
                     /function="involved in spore cortex synthesis (stage V
                     sporulation)"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAB12779.1"
                     /db_xref="GI:2633275"
                     /db_xref="GOA:P37875"
                     /db_xref="UniProt/Swiss-Prot:P37875"
                     /translation="MNAEEKKGLQRAIEEITEIAKGFGLDFYPMRYEICPAEIIYTFG
                     AYGMPTRFSHWSFGKQFHKMKLHYDLGLSKIYELVINSDPCYAFLLDSNSLIQNKLIV
                     AHVLAHCDFFKNNCRFQNTKRDMVESMAATAERIKHYEQVHGIKEVESFLDAILSIQE
                     HIDPSLVRPKLQWSVDDEEEEIEEAATPYDDLWSLDEKKPKKQVKKSKKPFPPRPEKD
                     ILLFIEEHSRELEPWQRDILTMMREEMLYFWPQLETKIMNEGWASYWHQRIIRELDLT
                     SDEAIEFAKLNAGVVQPSKTGINPYYLGLKIFEDIEKRYNNPTEEMKKMGVQPDSGRE
                     KMFEVREIESDISFIRNYLTKELVMREDLYLFQKQGRDYKIIDKQWKSVRDQLVSMRV
                     NGGFPYLTVNDGDYMKNNELYIKHWYEGIELDLKYLEKVLPYLFQLWGRSVHIESVLE
                     DKEVMFSYDGKGVHRRYL"
     gene            complement(5325. .6710)
                     /gene="phoA"
                     /locus_tag="BSU09410"
     CDS             complement(5325. .6710)
                     /gene="phoA"
                     /locus_tag="BSU09410"
                     /EC_number="3.1.3.1"
                     /note="alternate gene name: phoAIV"
                     /codon_start=1
                     /transl_table=11
                     /product="alkaline phosphatase A"
                     /protein_id="CAB12780.1"
                     /db_xref="GI:2633276"
                     /db_xref="GOA:P19406"
                     /db_xref="UniProt/Swiss-Prot:P19406"
                     /translation="MKKMSLFQNMKSKLLPIAAVSVLTAGIFAGAELQQTEKASAKKQ
                     DKAEIRNVIVMIGDGMGTPYIRAYRSMKNNGDTPNNPKLTEFDRNLTGMMMTHPDDPD
                     YNITDSAAAGTALATGVKTYNNAIGVDKNGKKVKSVLEEAKQQGKSTGLVATSEINHA
                     TPAAYGAHNESRKNMDQIANSYMDDKIKGKHKIDVLLGGGKSYFNRKNRNLTKEFKQA
                     GYSYVTTKQALKKNKDQQVLGLFADGGLAKALDRDSKTPSLKDMTVSAIDRLNQNKKG
                     FFLMVEGSQIDWAAHDNDTVGAMSEVKDFEQAYKAAIEFAKKDKHTLVIATADHTTGG
                     FTIGANGEKNWHAEPILSAKKTPEFMAKKISEGKPVKDVLARYANLKVTSEEIKSVEA
                     AAQADKSKGASKAIIKIFNTRSNSGWTSTDHTGEEVPVYAYGPGKEKFRGLINNTDQA
                     NIIFKILKTGK"
     gene            7242. .8273
                     /gene="lytE"
                     /locus_tag="BSU09420"
     CDS             7242. .8273
                     /gene="lytE"
                     /locus_tag="BSU09420"
                     /function="cell wall lytic activity"
                     /note="alternate gene name: papQ, cwlF"
                     /codon_start=1
                     /transl_table=11
                     /product="cell wall hydrolase; phosphatase-associated
                     protein (major autolysin) (CWBP33)"
                     /protein_id="CAB12781.1"
                     /db_xref="GI:2633277"
                     /db_xref="GOA:P54421"
                     /db_xref="UniProt/Swiss-Prot:P54421"
                     /translation="MKKQIITATTAVVLGALFAHTSIRELKEHIVRVLIRMRAMYKIS
                     EKRQYGLKKDKISKIKSENHLRSDIIYVGQTLSINGKSTSSKSSSSSSSSSTYKVKSG
                     DSFENFKKIGMTINELKKLNGLKSDLLRVGQVLKLKGSTSSSTPAHQKCASSSTSTYK
                     VKRETGLSKIARKYGTTVSKLKRLKRALKSDVIYVNQVLKVKGTSTSSSKPASSSSSS
                     SSKTSSTSLNVSKLVSDAKALVGTPYKWGGTTTSGFDCSGFIWYVLNKQTSVGRTSTA
                     GYWSSMKSIASPSVGDFVFFTTYKSGPSHMGIYIGNNSFIHAGSDGVQISSLNNSYWK
                     PRYLGAKRF"
     gene            complement(8292. .9218)

  Query Match             16.5%;  Score 166.6;  DB 1;  Length 198743;
  Best Local Similarity   51.5%;  Pred. No. 1.2e-16;
  Matches  460;  Conservative    0;  Mismatches  424;  Indels    9;  Gaps    3;

Qy     117 AAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAA 176
           | || ||  ||| |   | |  || ||   ||| |  || || || ||| | || || ||
Db   96088 AGAACATAAAGTGGCGGTAACACATGATTTAGGGAAGACAAATGTGCCTGAGCATCCGAA 96147

Qy     177 ACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACC 236
            || |||||||||||||||      ||| ||||| |  |  ||| | | |   |||  ||
Db   96148 GCGGGTTGTTGTTCTTGAGCTAGGTTTTATTGATACACTGCTTGATCTCGGCATTACGCC 96207

Qy     237 TGTTGGGATAGCGGATGATAACAAA---AAAAATCGTATTATTAAACCATTAAGAGATAA 293
           ||| ||| | || ||||| ||||||   ||  | |  || |  ||     |     | ||
Db   96208 TGTCGGGGTTGCCGATGACAACAAAGCGAAGCAGCTGATCAACAAGGATGTGCTGAAGAA 96267

Qy     294 AATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAA 353
           |||||     ||||| ||||| || || ||    || || | | | ||| |||||  |
Db   96268 AATTGACGGCTACACATCTGTCGGCACTCGCTCACAGCCAAGCATGGAAAAAATCGCTTC 96327
```

```
Qy     354 ACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTT 413
           | | ||||| ||||||||||||||||| |  |   | || ||     | ||  |  |  |
Db   96328 ATTAAAACCCGATTTAATTATTGCTGACACGACCCGGCATAAGAAGGTGTACGATCAGCT 96387

Qy     414 AAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATAT 473
            || ||||| || || ||||||| ||| || | |||  |||  |||||| | || |  ||
Db   96388 GAAAAAAATAGCGCCGACGATTGCACTTAATAATTTAAATGCTGATTATCAGGATACAAT 96447

Qy     474 TGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTT 533
           ||| ||||     || ||| |||||||  | || || ||  | |||    | |||    |
Db   96448 TGACGCTTCGCTTACGATTGCAAAAGCAGTCGGCAAGGAGAAGGAAATGGAGAAAAAGCT 96507

Qy     534 AGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCA 593
               |  || ||  | ||  ||   ||    ||  |  |||| |    | | |   |
Db   96508 GACGGCGCATGAAGAAAAGCTTAGCGAGACAAAGCAGAAAATCAGCGCGAACAGCCAGTC 96567

Qy     594 AAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTA 653
              | | |||   | |    |   | |  |   | ||       ||  || | || |
Db   96568 CGTGCTTTTGATCGGAAATACAAATGATACCATTATGGCCAGGGATGAAAACTTCTTTAC 96627

Qy     654 TGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAA 713
                | |  |   |||  ||  | || |    | | ||| | |||      |  | ||
Db   96628 ATCGAGACTTT---TAACACAGGTCGGCTACCGATATGCAATCAGTACGTCAGGCAATAG 96684

Qy     714 AGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCA 773
            | || ||| ||        |        |  |  ||||||     ||     |     |
Db   96685 CGATTCAAGCAATGGCGGCGACTCTGTGAATATGAAAATGACACTGGAGCAGCTGCTGAA 96744

Qy     774 AGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACT 833
           |    |||| ||    ||  || | ||||||   ||| ||  |     |||  |
Db   96745 AACAGATCCGGATGTGATCATCCTGATGACAGGAAAAACAGATGACCTCGACGCCGACGG 96804

Qy     834 AAAAGAAC---TAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACG 890
            |||   |   | |||||  ||    | ||||| ||| | || || ||||||||    |
Db   96805 TAAACGCCCGATCGAAAAGAATGTCCTTTGGAAAAAACTGAAGGCAGTGAAAAACGGGCA 96864

Qy     891 TGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGA 950
           |||  |     | || ||||   | ||| |       ||  || ||       |  | |
Db   96865 TGTATACCACGTGGATCGTGCGGTGTGGTCTCTGCGCCGCAGTGTAGACGGGGCGAATGC 96924

Qy     951 AATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATA 1003
            ||     | || |||    ||        |   | ||  ||| || | || |
Db   96925 CATTTTGGACGAGCTTCAAAAAGAGATGCCGGCTGCTAAGAAATAAAAGAAAA 96977


RESULT 25
AR356438
LOCUS       AR356438                 242 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 2556 from patent US 6593114.
ACCESSION   AR356438
VERSION     AR356438.1  GI:33762522
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 242)
  AUTHORS   Kunsch,C.A., Choi,G.H., Barash,S., Dillon,P.J., Fannon,M.R. and
            Rosen,C.A.
  TITLE     Staphylococcus aureus polynucleotides and sequences
  JOURNAL   Patent: US 6593114-A 2556 15-JUL-2003;
FEATURES             Location/Qualifiers
     source          1. .242
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             16.1%;  Score 162.2;  DB 6;  Length 242;
  Best Local Similarity   79.3%;  Pred. No. 3.2e-15;
  Matches  191;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy     208 GATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAAT 267
           ||||| ||||  || ||||| |||||||| ||||| || || ||||||   || |||||
Db       2 GATGCATTAGCAGCATTAGACGTTAAACCAGTTGGTATTGCTGATGATGGTAAGAAAAAA 61
```

```
Qy        268 CGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAG 327
              ||||| ||||||||| | ||||| ||||||||  | || ||||||||||| |||||||| 
Db         62 CGTATCATTAAACCAGTTAGAGAAAAAATTGGGGATTATACTTCTGTAGGTACACGTAAA 121

Qy        328 CAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAAT 387
              || || |||||||||||||| |||||| | ||||| ||||||||||| |||||||  | |
Db        122 CAGCCAAACTTAGAAGAAATTAGTAAATTAAAACCGGATTTAATTATCGCTGATAGCAGT 181

Qy        388 AGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGT 447
              ||||| | ||||||| ||||||| ||||  |||||||| ||  |   | | || |||
Db        182 AGACATANAGGTATTAATAAAGAATTAANCAAAATTGCACCAACATTATCATTAAAGAGT 241

Qy        448 T 448
              |
Db        242 T 242


RESULT 26
AR537994
LOCUS       AR537994                 242 bp    DNA     linear   PAT 08-OCT-2004
DEFINITION  Sequence 2556 from patent US 6737248.
ACCESSION   AR537994
VERSION     AR537994.1  GI:53929211
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 242)
  AUTHORS   Kunsch,C.A., Choi,G.A., Barash,S.C., Dillon,P.J., Fannon,M.R. and
            Rosen,C.A.
  TITLE     Staphylococcus aureus polynucleotides and sequences
  JOURNAL   Patent: US 6737248-A 2556 18-MAY-2004;
FEATURES             Location/Qualifiers
     source          1..242
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             16.1%;  Score 162.2;  DB 6;  Length 242;
  Best Local Similarity   79.3%;  Pred. No. 3.2e-15;
  Matches  191;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy        208 GATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAAT 267
              ||||| ||||  || ||||| |||||||| ||||| || || ||||||   || ||||| 
Db          2 GATGCATTAGCAGCATTAGACGTTAAACCAGTTGGTATTGCTGATGATGGTAAGAAAAAA 61

Qy        268 CGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAG 327
              ||||| ||||||||| | ||||| ||||||||  | || ||||||||||| |||||||| 
Db         62 CGTATCATTAAACCAGTTAGAGAAAAAATTGGGGATTATACTTCTGTAGGTACACGTAAA 121

Qy        328 CAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAAT 387
              || || |||||||||||||| |||||| | ||||| ||||||||||| |||||||  | |
Db        122 CAGCCAAACTTAGAAGAAATTAGTAAATTAAAACCGGATTTAATTATCGCTGATAGCAGT 181

Qy        388 AGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGT 447
              ||||| | ||||||| ||||||| ||||  |||||||| ||  |   | | || |||
Db        182 AGACATANAGGTATTAATAAAGAATTAANCAAAATTGCACCAACATTATCATTAAAGAGT 241

Qy        448 T 448
              |
Db        242 T 242


RESULT 27
AE006047/c
LOCUS       AE006047               13792 bp    DNA     linear   BCT 16-JUN-2004
DEFINITION  Pasteurella multocida subsp. multocida str. Pm70 section 14 of 204
            of the complete genome.
ACCESSION   AE006047 AE004439
VERSION     AE006047.1  GI:12720328
KEYWORDS    .
SOURCE      Pasteurella multocida subsp. multocida str. Pm70
  ORGANISM  Pasteurella multocida subsp. multocida str. Pm70
            Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
            Pasteurellaceae; Pasteurella.
REFERENCE   1  (bases 1 to 13792)
  AUTHORS   May,B.J., Zhang,Q., Li,L.L., Paustian,M.L., Whittam,T.S. and
            Kapur,V.
  TITLE     Complete genomic sequence of Pasteurella multocida, Pm70
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)
  MEDLINE   21145866
   PUBMED   11248100
REFERENCE   2  (bases 1 to 13792)
  AUTHORS   Zhang,Q. and Kapur,V.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-OCT-2000) Department of Veterinary Pathobiology,
            University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN
            55108, USA
FEATURES             Location/Qualifiers
     source          1..13792
                     /organism="Pasteurella multocida subsp. multocida str.
                     Pm70"
                     /mol_type="genomic DNA"
                     /strain="PM70"
                     /sub_species="multocida"
                     /db_xref="taxon:272843"
     gene            87..809
                     /gene="PM0116"
     CDS             87..809
                     /gene="PM0116"
                     /codon_start=1
                     /transl_table=11
                     /product="unknown"
                     /protein_id="AAK02200.1"
                     /db_xref="GI:12720329"
                     /translation="MTTFIAYGNINDTYPLDSIPRELLIEKLRQPQSSNQRVQKRHRC
                     RWVAHFLLWQLLKISQKPTALLKHIDYSQSGRPQLPVHDVDFNISHSGDWVAVILRVN
                     TQGDSIVGIDIESPQKERDYPALLAYFASPSEQAWFTQQQEKEKSFYLSWCLREAVLK
                     SQGVGIVKLSAVQHDPLLRQIQSQHCPAGQLVFSHRLPFYLAFFINQTQCLDVHYFTW
                     QRGILAETTLHDAIYYKVNEKT"
     gene            994..2253
                     /gene="hflK"
                     /note="synonym: PM0117"
     CDS             994..2253
                     /gene="hflK"
                     /codon_start=1
                     /transl_table=11
                     /product="HflK"
                     /protein_id="AAK02201.1"
                     /db_xref="GI:12720330"
                     /translation="MENVMSLNGSDQDPWGKPGQKKPESQRPDNEGGQPNWNDNQSGR
                     KQNNQEQSPPDIEEIFNNLLKKLGGSGNGGKRNNSHSGGSLNLGKLLPIAAVIGAIVW
                     GVSGFYTVKEAERGVVMRFGELHAIVQPGLNWKPTFIDRVIPVNVEQVKELRTQGSML
                     TQDENMVKVEMTVQYRVHDPAKYLFSVTNADDSLNQATDSALRYVIGHMSMDDILTTG
                     RSVVRENTWKTLNTIIEPYNMGLEVVDVNFQSARPPEEVKDAFDDAIKAQEDEQRYIR
                     EAEAYAREREPIARGDAQRILEEATAYKDRVVLDAKGEVERFERLLPEFKAAPELLRE
                     RLYIQTMEKVMANTPKVMLDGNSGNNLTVLPLEQILKGQQTKGKSTNATSEEVRPVFR
                     AETPNVMPTQTSTSSTRQENTRQGRFN"
     gene            2253..3140
                     /gene="hflC"
                     /note="synonym: PM0118"
     CDS             2253..3140
                     /gene="hflC"
                     /codon_start=1
                     /transl_table=11
                     /product="HflC"
                     /protein_id="AAK02202.1"
                     /db_xref="GI:12720331"
                     /translation="MRKFLTPVIVVIAAILYSSIVIVSEGTRGIMLRFSKVHRDADNK
                     VVVYNPGLHFKIPFIDSIKILDARIRTLDGQADRFVTVEKKDLLVDSYVKWRISDFGR
                     FYTATGGGDYAQASNLLRRKVNDRLRSETGSRTIKDIVSGTRGELMEGARKALNTGPD
                     STAELGIEVVDVRVKQINLPDEVSSSIYQRMRAERDAVAREHRSQGREKAAFIQADVD
                     RKVTLILANANRTAQELRGSGDATAAKVFSDAFSQEPQFYSFLRSLKAYESSFANSDN
                     MMILKPDSDFFRFMQAPTK"
     gene            complement(3197..3739)
```

```
                     /gene="PM0119"
     CDS             complement(3197. .3739)
                     /gene="PM0119"
                     /codon_start=1
                     /transl_table=11
                     /product="unknown"
                     /protein_id="AAK02203.1"
                     /db_xref="GI:12720332"
                     /translation="MAKRGKNKIEWDNTDWEQEEEEIIWVSKSEIKRDAEALKKLGEK
                     LVDLTKAKLEKIPLEEKLLEAIELAQRLQKEARRRQLQYIGKLLRNIDVEPIQDALEK
                     LENKHQQQQAMLHKLEMLRDTLIDNPEAITQFLEQHPDADRQHLRNLVRAAQKEKAQN
                     KPPKAYREIFQYLKDFMLED"
     gene            3834. .5189
                     /gene="pmbA"
                     /note="synonym: PM0120"
     CDS             3834. .5189
                     /gene="pmbA"
                     /codon_start=1
                     /transl_table=11
                     /product="PmbA"
                     /protein_id="AAK02204.1"
                     /db_xref="GI:12720333"
                     /translation="MQNTEKSTALLKQQEQTLREAVRYAVDLAQKAGATAEVAVTKVS
                     GLSVSTRLQEIENVEFTNDGALGISVYLGNQKGNASTSDLSPEAIKNTVEAAISIAKY
                     ASPDECAGLADYDLMAFDAPDLALYHAADIDVEKATALALETEQAALSYDKRIVNSEG
                     SAFNSHTGVRVYGNTHGMLQSYLSSRYSLSCSVIAGVEDQLERDYEYTISRDLNQLES
                     ASWVGEHCAKKTLSRLQPQKMATQQVPVIFLNDVATGLISHLTAAISGGSLYRKSSFL
                     LDHLGKQVLPDWFHISEQPHLIGQLASSPFDSEGVRTQCREIIQQGTLQTYLLTSYSG
                     KKLGMRSTGHAGGIHNWLVKPNTTGGLTTLLRQMGTGLLVTDLMGQGVNLVTGDYSRG
                     AAGFWVENGEIQYPVAEITIAGQLKEMLRNIVAVGDDIEHRSNIQTGSILLEHLKVSG
                     N"
     gene            5424. .5963
                     /gene="hpt"
                     /note="synonym: PM0121"
     CDS             5424. .5963
                     /gene="hpt"
                     /codon_start=1
                     /transl_table=11
                     /product="Hpt"
                     /protein_id="AAK02205.1"
                     /db_xref="GI:12720334"
                     /translation="MKKHHVGILISEQDVRARIQTLGREITQYYQQKAVEKVIVVGLL
                     RGSFMFMADLVRELNLPVEIEFMTTSSYGSGMTTNHDVKISKDLDGDIKGEHVLIIED
                     IIDTGYTLQKVREILKLREPASLRICTLLDKPSRREVEVPVEWVGFSIPDEFVVGYGI
                     DYAQRYRNLGYIGHVVLEE"
     gene            5996. .6580
                     /gene="gmhA"
                     /note="synonym: PM0122"
     CDS             5996. .6580
                     /gene="gmhA"
                     /codon_start=1
                     /transl_table=11
                     /product="GmhA"
                     /protein_id="AAK02206.1"
                     /db_xref="GI:12720335"
                     /translation="MYLEQIKAELNEAATVLNKFINDENNMKLIQQAALLISDSFKQG
                     GKVLSCGNGGSHCDAMHFAEELTGRYRENRPGYPAIAISDVSHLSCVSNDFGYEYVFS
                     RYVEAVGQKGDVLFGLSTSGNSKNILNAIKAAKEKGMKVIAMTGKDGGQMAGIADVEI
                     RVPHFGYADRIQEIHIKVIHILMMLIEFEMAKTE"
     gene            6756. .7487
                     /gene="artP"
                     /note="synonym: PM0123"
     CDS             6756. .7487
                     /gene="artP"
                     /codon_start=1
                     /transl_table=11
                     /product="ArtP"
                     /protein_id="AAK02207.1"
                     /db_xref="GI:12720336"
                     /translation="MAIRVKNLNFFYGSSQALFDINLSAEEGDTVVLLGPSGAGKSTL
                     IRTLNLLEVPASGELIIANSQFDLSETKTNPKQICRLRQDVGMVFQQYHLWPHLTVIE
                     NLIEAPKKVLKLSDEQARSEALELLKRLRLAQFADRFPLHLSGGQQQRVAIARALMMK
                     PQVLLFDEPTAALDPEITAQVVDIIKELQQTGITQVIVTHEVGVAKKVATKVVYMEQG
```

```
                     HIIEMGDASCFTNPKTEQFKQYLSH"
     gene            7504. .8226
                     /gene="artI"
                     /note="synonym: PM0124"
     CDS             7504. .8226
                     /gene="artI"
                     /codon_start=1
                     /transl_table=11
                     /product="ArtI"
                     /protein_id="AAK02208.1"
                     /db_xref="GI:12720337"
                     /translation="MKKLLLTALLASTALAANAKEITFAMEPSYPPFELTNEKGEIIG
                     FDVDIANAICKEIQATCHFKGQAFDSLITGLRSKHFDAAISAIDITEARAKQVSFSDP
                     YYDSSSSFISVKGKVDDLAQAKKVGVQNGTTYQQYIVAEAKQYQPVPYASLQNAILDL
                     KNGRIDVIFGDTAVLVDMLAKEQSLAFVGDKVTNKKYFGNGFGIAVNKASKDLLDDLN
                     KGLATIKANGEYQKIYDKWMTK"
     gene            8233. .8904
                     /gene="artQ"
                     /note="synonym: PM0125"
     CDS             8233. .8904
                     /gene="artQ"
                     /codon_start=1
                     /transl_table=11
                     /product="ArtQ"
                     /protein_id="AAK02209.1"
                     /db_xref="GI:12720338"
                     /translation="MFFDYLSLMYTATLMTLGLAVCSLIVGLILSIIFVVLETSKWQW
                     IRKPTSVFVTLLRGLPEILVVLLIYFGSTEVVEMLTGEFIQFSAFSCGVFALSIIFAA
                     YASQSLRGAIQAIPLGQWESGAALGLSRGYTFVQIIMPQVWRHALPGLSNQWLVLLKD
                     TALVSLIGVEDLMRQAELVNTVTHQPFTWYGIAALLYLTITLISQFIIRKLEFRFTRF
                     ERGVE"
     gene            8904. .9587
                     /gene="artM"
                     /note="synonym: PM0126"
     CDS             8904. .9587
                     /gene="artM"
                     /codon_start=1
                     /transl_table=11
                     /product="ArtM"
                     /protein_id="AAK02210.1"
                     /db_xref="GI:12720339"
                     /translation="MFQDYFMVIAQGIPTSLLLTFVSLVVAFFLALLLTFGLSIGNKL
                     VKCGIHTYLTLFTGTPLLVQFFLIYAGPGQFEWLVDSALWPLLSNAWFCAMLTLALNS
                     AAYSTLLFHGAVKAISKGQWESCAALGLTRMETLKILIPYALKRALPSYSNEIILVFK
                     GTALASTITIMDIMGYARQLYGTEYDALTIYGIAGMIYLIITAIATLLLRKLESKVLS
                     FERLEVEKA"
     gene            complement(9646. .9921)

  Query Match             15.5%;  Score 156.6;  DB 1;  Length 13792;
  Best Local Similarity   51.9%;  Pred. No. 8.1e-15;
  Matches  456;  Conservative    0;  Mismatches  384;  Indels   39;  Gaps    3;

Qy        123 TGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGT 182
              ||  |||    ||||  |  || ||||       |   ||  ||     |||||||||||
Db      13549 TGCTGTTACCGTCAAAGATCAAAAAGGGGAATTTACCTTAGATAGCGTACCTAAACGTGT 13490

Qy        183 TGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGG 242
              |||||   | || ||||| | ||| |||||  | |      | | ||| |  ||||||||
Db      13489 TGTTGCCTTAGAATATTCTTATGTGGATGCACTTGCACAAATTGGTGTCAGCCCTGTTGG 13430

Qy        243 GATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAA 302
                | || ||||| ||  | ||||  |||||| |  ||  | || | |||||| |
Db      13429 TGTCGCAGATGACAATGATAAAACCCGTATTCTGCAAAAAGTACGCGATAAAGTGCAGCC 13370

Qy        303 ATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACC 362
              ||     || || || || |||   ||||| |  |||||||  ||   |  |||||||||
Db      13369 ATGGGAATCAGTGGGGACGCGTTCTCAACCGAGTTTAGAAGCGATTTCTGCACTTAAACC 13310

Qy        363 AGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAAT 422
              |||||| || ||||| ||| ||||| | ||    |   | ||| ||||  | || |||||
Db      13309 AGATTTGATCATTGCCGATGATAATCGCCATTCTGCCGTCTATGAAGAACTCAAAAAAAT 13250

Qy        423 TGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTT 482
```

```
               || || ||  | |   | ||      | |||   | ||| | |||||  ||||  |
Db      13249 CGCGCCGACAGTCGTCTTTAATTCCCGCCATGAGAACTATCAAGAAAACCTTGAAACCGC 13190

Qy        483 TAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACA 542
                ||| |||     |    |||||||||  | |||||    ||   ||  | | | ||||
Db      13189 ACAAAAAATCGGTGATTTATTAGGTAAATCAAAAGAAATGCAAGCGCGTATTGCAAAACA 13130

Qy        543 CGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATT 602
                |  || | ||||  ||       |  |  | |||  |    |||    |||| |   |
Db      13129 TAAGCAGGATATTGCGGAC------ATCGCCAAAACGTTACCGAAAGGAAAAAAAGCGAT 13076

Qy        603 GCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCA 662
                  |  ||  |   | || |    ||   |    ||   ||| | || ||||| ||||
Db      13075 TATCGGTGTTTCACGTGAAACCCAATTTAATTTATATAATAGCGAATCCTATGCTGGTGG 13016

Qy        663 ATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAG 722
               ||  |       ||||||| | ||    ||  |     ||  | ||   ||  |
Db      13015 CTTAGTGGAAGTGCTAGGTTATCAAATGCCAAAAGCCCGTGCCGATAACCAACCTAATGC 12956

Qy        723 TAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCC 782
              |  |  |  |  || ||                            |  |   || || ||
Db      12955 TTCGGTTGGTTTAGAACAA------------------------GTGGCGGCAGAAAAGCC 12920

Qy        783 TGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACT 842
              ||| |  ||| || ||||     |         |  | | ||   |     || | ||
Db      12919 TGATCTGATGATCTTAATCCATTA---------TCGTGATGAGAGTATTGCAAGAAAATG 12869

Qy        843 AGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTT 902
               ||||| ||  |  |||||||| |   |    || || ||||||   |  ||     |
Db      12868 GGAAAATGAAGCCTTATGGAAAATTATCCCTGCGGTAAAAAATGGTCAAGTCATCTTAGC 12809

Qy        903 AGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGA 962
                |   | | |||||||||||| |||||||| |   | |    | |||| ||||||| |||||
Db      12808 TAATGATAATTTATGGGCAAGAGCACGTGGTATTGATGCCGCTGAAGTAATGGCTAAAGA 12749

Qy        963 ACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAA 1001
              | |    || ||    | ||||      ||| ||   ||
Db      12748 AGTCCAAGACTTTGTAACGAAATCCGCCAAATAATGAAA 12710
```

```
RESULT 28
AX433651
LOCUS       AX433651                 801 bp    DNA     linear   PAT 28-JUN-2002
DEFINITION  Sequence 2066 from Patent WO0229113.
ACCESSION   AX433651
VERSION     AX433651.1  GI:21658459
KEYWORDS    .
SOURCE      Bacillus licheniformis
  ORGANISM  Bacillus licheniformis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE   1
  AUTHORS   Berka,R. and Clausen,I.G.
  TITLE     Methods for monitoring multiple gene expression
  JOURNAL   Patent: WO 0229113-A 2066 11-APR-2002;
            Novozymes Biotech, Inc. (US) ; Novozymes A/S (DK)
FEATURES             Location/Qualifiers
     source          1..801
                     /organism="Bacillus licheniformis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1402"
ORIGIN

  Query Match             14.8%;  Score 149.4;  DB 6;  Length 801;
  Best Local Similarity   51.8%;  Pred. No. 2.1e-13;
  Matches  414;  Conservative    0;  Mismatches  376;  Indels    9;  Gaps    3;

Qy        176 AACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAAC 235
              ||   ||||| |||||||| |    |||| |||||||| |  ||| |   | | |||| |
Db          1 AAAAAGTTGTGGTTCTTGAATTGGGATTTATTGATGCGCTCCTTGATGCGGGTATTAAGC 60

Qy        236 CTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAA 295
              |||| || || ||||| ||   ||||   ||   |||||    |     | || |  ||||
Db         61 CTGTCGGAATTGCGGACGACGGCAAACCTAAGTTTATTAACGAGAAGGTCAGGGGAAAAA 120
```

```
Qy        296 TTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAAC 355
              |   |   || ||||| || ||  | ||    || || | ||| ||| ||||   |
Db        121 TCAAAGGGTATACTTCAGTCGGTTCGCGCGCCCAGCCAAGCTTTGAAAAAATTGCTTCTT 180

Qy        356 TTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAA 415
              | || || |||||||||||||| |||   |  || ||||  ||  | ||| |  |  |
Db        181 TAAAGCCCGATTTAATTATTGCCGATTCGAGCAGGCACAGCGGCGTCTATGATAAGCTGT 240

Qy        416 ATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTG 475
                |||||||| || || || |  || ||||  ||  |||  |||||| | || |   |||
Db        241 CGAAAATTGCGCCGACAATCGCGCTCAAAAACTTGAATGCCGATTATCAGGACACGCTTG 300

Qy        476 ATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAG 535
              |||| | |   || ||| | |||||  | || ||||||   | |    | |||   || |
Db        301 ATGCATCTCTTACGATTGCGAAAGCCGTCGGCAAAGAAAGCGCAATGGAGAAAAAATTGG 360

Qy        536 AAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAA 595
                |||||  |  | ||  | || ||||  ||| |  || |    |     ||| |    |
Db        361 CTGAACATAAACAAAAGCTGGATGAATTGAAACAGAAATTCGGCAGCCGGAAACAGAGCA 420

Qy        596 AGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATG 655
                 |  |||  |      |   | ||  |       |     ||  |   |||| ||
Db        421 TTCTTCTGCTCGGGAACACAAATGAAGAAATCACCGTGCGCGATGAA---AACTTTTTCA 477

Qy        656 TTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAG 715
                  |||  | || |   || | || | || | | |   |  | ||    |     ||||
Db        478 CGTCTCAGCTTCTGACGAAAATCGGCTATACATACGGTGTCGGAGACAGCG---GCAAAG 534

Qy        716 GTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAG 775
              |        |  |   | ||       |  |  ||||||     |||     |   | |
Db        535 GCGATGCCGAAAACGGTGAATCCGTCAATATTAAAATGACGCTCGAACAGCTGCTTGAGA 594

Qy        776 TGAATCCTGAGCGTATGTTCATAATGAC---AAACAAAGCAAGTTCTAACGAACCTTCAC 832
                 |||| ||   |||  || | |||||     | | |||| |      |   |||
Db        595 AAGATCCCGACGTTATTGTCCTGATGACGGGAGAAAAAGATAAAGTCGATGAAGACGGGA 654

Qy        833 TAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTG 892
               || |    | ||||||||||||| | |||||| ||  | | ||| || |||||       |
Db        655 AAAGACCGATCGAAAAAGATCCTCTTTGGAAAAAGCTCAGCGCAGTCAAAAACGGCAAGG 714

Qy        893 TTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAA 952
              || ||      ||| |   |   ||| |      |||  |  |        |||| ||
Db        715 TTTATGAAGCCGACAGATTCGCCTGGTCGCTCCGACGCAGCATTGACGGAGCAGATGAGC 774

Qy        953 TGGCAAAAGAACTTGTTGA 971
              ||    | ||| | | | |
Db        775 TGATGGATGAAATCGATCA 793
```

```
RESULT 29
AE017012
LOCUS       AE017012               301332 bp    DNA     linear   BCT 16-MAY-2003
DEFINITION  Bacillus cereus ATCC 14579 section 15 of 18 of the complete genome.
ACCESSION   AE017012 AE016877
VERSION     AE017012.1  GI:29897926
KEYWORDS    .
SOURCE      Bacillus cereus ATCC 14579
  ORGANISM  Bacillus cereus ATCC 14579
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1  (bases 1 to 301332)
  AUTHORS   Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelon,B.,
            Kapatral,V., Bhattacharyya,A., Reznik,G., Mikhailova,N.,
            Lapidus,A., Chu,L., Mazur,M., Goltsman,E., Larsen,N., D'Souza,M.,
            Walunas,T., Grechkin,Y., Pusch,G., Haselkorn,R., Fonstein,M.,
            Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.
  TITLE     Genome sequence of Bacillus cereus and comparative analysis with
            Bacillus anthracis
  JOURNAL   Nature 423 (6935), 87-91 (2003)
   MEDLINE  22608415
    PUBMED  12721630
REFERENCE   2  (bases 1 to 301332)
```

```
  AUTHORS   Candelon,B., Gailloux,K., Ehrlich,D.S. and Sorokin,A.
  TITLE     The number of ribosomal RNA operons in Bacillus cereus
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 301332)
  AUTHORS   Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelon,B.,
            Kapatral,V., Bhattacharyya,A., Reznik,G., Mikhailova,N.,
            Lapidus,A., Chu,L., Mazur,M., Goltsman,E., Larsen,N., D'Souza,M.,
            Walunas,T., Grechkin,Y., Pusch,G., Haselkorn,R., Fonstein,M.,
            Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAR-2003) INRA, Genetique Microbienne, Domaine de
            Vilvert, Jouy en Josas 78352, France
FEATURES             Location/Qualifiers
     source          1. .301332
                     /organism="Bacillus cereus ATCC 14579"
                     /mol_type="genomic DNA"
                     /strain="ATCC 14579"
                     /db_xref="ATCC:14579"
                     /db_xref="taxon:226900"
     gene            complement(114. .470)
                     /locus_tag="BC4288"
     CDS             complement(114. .470)
                     /locus_tag="BC4288"
                     /codon_start=1
                     /transl_table=11
                     /product="Membrane-attached cytochrome c550"
                     /protein_id="AAP11201.1"
                     /db_xref="GI:29897927"
                     /translation="MKRNPLIPFALIAALGIIVMFVFSFQGLNKSKELADAKNGGKPA
                     QTASKPEDIVKQSCTSCHGDQLQGAVGPNLQKIGGKLSKDEIKEILSKGKGNMPANIV
                     PADQAAKVADWLSKKK"
     gene            complement(857. .1984)
                     /locus_tag="BC4289"
     CDS             complement(857. .1984)
                     /locus_tag="BC4289"
                     /codon_start=1
                     /transl_table=11
                     /product="RNA polymerase sigma factor rpoD"
                     /protein_id="AAP11202.1"
                     /db_xref="GI:29897928"
                     /translation="MADKPARSKQIETEMTLEQVKEQLTELGKKRGVLTYEEIAERMN
                     GFEIESDQMDEYYEYLGEQGIDLVGDNDNEEGPNNRQISKTEEEFDLNDLSVPPGVKI
                     NDPVRMYLKEIGRVDLLSAEEEIRLATRIEEGDEEAKRRLAEANLRLVVSIAKRYVGR
                     GMLFLDLIQEGNMGLIKAVEKFDYRKGFKFSTYATWWIRQAITRAIADQARTIRIPVH
                     MVETINKLIRVQRQLLQDLGREPSPEEIGEEMDLAPEKVREILKIAQEPVSLETPIGE
                     EDDSHLGDFIEDQEATSPADHAAYELLKEQLEDVLDTLTDREENVLRLRFGLDDGRTR
                     TLEEVGKVFGVTRERIRQIEAKALRKLRHPSRSKRLKDFLE"
     gene            complement(2044. .3840)
                     /locus_tag="BC4290"
     CDS             complement(2044. .3840)
                     /locus_tag="BC4290"
                     /EC_number="2.7.7.-"
                     /codon_start=1
                     /transl_table=11
                     /product="DNA primase"
                     /protein_id="AAP11203.1"
                     /db_xref="GI:29897929"
                     /translation="MGNRIPEEVVEQIRTSSDIVEVIGEYVQLRKQGRNYFGLCPFHG
                     ENSPSFSVSSDKQIFHCFGCGEGGNVFSFLMKMEGLAFTEAVQKLGERNGIAVAEYTS
                     GQGQQEDISDDTVIMQQAHELLKKYYHHLLVNTEEGNEALSYLLNRGITKEMIEKFEI
                     GYASPAWDAATKILQKRGLSLSTMEQAGLLIRSEKDGSHYDRFRGRVMFPIYTLQGKV
                     IAFSGRALGDDTPKYLNSPETPIFHKSKLLYNFHQARPFIRKRGQVVLFEGYADVLAA
                     VKSGVEEAVATMGTALTEEQAKLLRRNVETVVLCYDGDKAGREATMKAGQLLLKVGCQ
                     VKVTSLPDKLDPDEYVQQYGTTAFENLVKSSVSFVGFKINYLRLGKNLQDESGKEEYV
                     KSVLKELSLLQDAMQAESYLKSLSQEFSYSMETLLNQLHQYRKEQKVQQKQIKQVAKP
                     SQVVQTKPKLTGFERAEREIIYHMLQSAEVAVRMESHIEDFHTEEHKGILYELYAYYE
                     KGNEPSVGTFLSWLSDEKLKNIITDISTDEFINPEYTEEVLQGHLETLRRHQEKLEKM
                     EIIFKVKQMEKTDPVEAAKYYVAYLQNQKARK"
     gene            4007. .4213
                     /locus_tag="BC4291"
     CDS             4007. .4213
                     /locus_tag="BC4291"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAP11204.1"
                     /db_xref="GI:29897930"
                     /translation="MLYKIVYICSSVTNVSLRNNTYTPFYKKSSKPHILFDLEICNFF
                     MLIHITINSILFSIYKDPIALLFP"
     gene            complement(4202. .4333)
                     /locus_tag="BC4292"
     CDS             complement(4202. .4333)
                     /locus_tag="BC4292"
                     /codon_start=1
                     /transl_table=11
                     /product="EMG2 protein"
                     /protein_id="AAP11205.1"
                     /db_xref="GI:29897931"
                     /translation="MQNINNILVGVDVYPVKAKIVQVGAEFHVGILFVASYAHPLRK"
     gene            complement(4701. .5513)
                     /locus_tag="BC4293"
     CDS             complement(4701. .5513)
                     /locus_tag="BC4293"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical Cytosolic Protein"
                     /protein_id="AAP11206.1"
                     /db_xref="GI:29897932"
                     /translation="MDNKIVYVVSDSVGETADLVVRAAMGQFPFAPDIRRVPYVEDTG
                     TLKEVISIAKSNQALICFTLVKPDMRQYLVTEAAKEGVEAYDIIGPLIDQIEEITGQV
                     PRYEPGVVRRLDEEYFKKIEAIEFAVKYDDGRDARGILKADIVLIGVSRTSKTPLSQY
                     LAHNKRLKVANVPLVPEVDPPEELYQVAKEKCFGLKITPDKLNHIRKERLKSLGLSDG
                     ATYANINRIQEEIDHFEEVISRINCQVIDVSNKAIEETANIIVNAVQNQKMF"
     gene            complement(5543. .6175)
                     /locus_tag="BC4294"
     CDS             complement(5543. .6175)
                     /locus_tag="BC4294"
                     /codon_start=1
                     /transl_table=11
                     /product="CBS domain containing protein"
                     /protein_id="AAP11207.1"
                     /db_xref="GI:29897933"
                     /translation="MIIIELNKRQEHIIQIVKDHGPITGESIAAQLGLTRATLRPDLA
                     ILTMAGYLEARPRVGYFYTGKTGGQLLSEAVKKIKVQDYQSRPVVIDKNVSVYDAICT
                     MFLEDVGTLFVVDQMTLLVGVVSRKDLLRASLGKQDLTSLPVNIIMTRMPNIAMCRRE
                     DSLYDIAMELIERQIDAMPVVKDTKQGLEVIGRITKTNITRAFVNLVNNE"
     gene            complement(6283. .7029)
                     /locus_tag="BC4295"
     CDS             complement(6283. .7029)
                     /locus_tag="BC4295"
                     /codon_start=1
                     /transl_table=11
                     /product="DNA repair protein recO"
                     /protein_id="AAP11208.1"
                     /db_xref="GI:29897934"
                     /translation="MFQKVEGIVIRTTDYGETNKIVTIFSRELGKVSAMARGSKKPKS
                     RLASVSQLMTHGHFLIQMGSGLGTLQQGEIISTMKEIREDIFLTAYASFIVELTDKAT
                     EDKKHNPYLFEMLYQTLHYMCEGVDPEVLSLIYQTKMLPVLGMRPYFDTCAICHQETD
                     FVAFSVREGGFLCSRHAEQDPYRIPVGEAVHKLLRLFFHFDLHRLGNVSVKDSTKKQM
                     RLVLNTYYDEYCGIYLKSRRFLEQLDKFQI"
     gene            complement(7063. .7206)
                     /locus_tag="BC4296"
     CDS             complement(7063. .7206)
                     /locus_tag="BC4296"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAP11209.1"
                     /db_xref="GI:29897935"
                     /translation="MLDFTWKFFSKTGSIETYLLLKEMEKDVNDEMDQHEEELAHLDS
                     PIS"
     gene            complement(7383. .8288)
                     /locus_tag="BC4297"
     CDS             complement(7383. .8288)
                     /locus_tag="BC4297"
                     /codon_start=1
                     /transl_table=11
                     /product="GTP-binding protein"
                     /protein_id="AAP11210.1"
                     /db_xref="GI:29897936"
                     /translation="MNRKGYKSGFVSIIGRPNVGKSTFLNRIIGQKIAIMSDKPQTTR
                     NKIQGVYTENDAQVIFIDTPGIHKPKHKLGDFMVKMAQTTLKEVDIVLFMVNAVEGFG
                     RGEEFIIEKLKETKQPVFLVINKIDQLHPEQLLELIDQYRKLHEFAEIVPISALDGNN
                     VDALIGTIKKYLPEGPQYYPENQVTDHPERFIIAELIREKVLHLTREEVPHSVAVVID
                     AIQKREGGAVYINATIVVERASQKGIIIGKQGKMLKEVGKRARFDIEALLGSKVFLEV
                     WVKVQKDWRNKMSQLRDLGFREDEY"
     gene            complement(8281. .8679)
                     /locus_tag="BC4298"
     CDS             complement(8281. .8679)
                     /locus_tag="BC4298"
                     /EC_number="3.5.4.5"
                     /codon_start=1
                     /transl_table=11
                     /product="Cytidine deaminase"
                     /protein_id="AAP11211.1"
                     /db_xref="GI:29897937"
                     /translation="MNSKQLILEAIEARKQAYVPYSKFQVGAALLTRGGKVYRGCNVE
                     NASYGLCNCAERTALFKAVSEGDKEFVAIAVVADTKRPVPPCGACRQVMVELCKQDTK
                     VYLSNLHGDVQETTVGELLPGAFLAEDLHE"
     gene            complement(8780. .9133)
                     /locus_tag="BC4299"
     CDS             complement(8780. .9133)
                     /locus_tag="BC4299"
                     /EC_number="2.7.1.107"
                     /codon_start=1
                     /transl_table=11
                     /product="Diacylglycerol kinase"
                     /protein_id="AAP11212.1"
                     /db_xref="GI:29897938"

  Query Match             14.5%;  Score 146.4;  DB 1;  Length 301332;
  Best Local Similarity   55.4%;  Pred. No. 1.3e-13;
  Matches  310;  Conservative    0;  Mismatches  241;  Indels    9;  Gaps    1;

Qy         25 AAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAAT 84
              |||||| | ||| |  |     || |  || | ||   ||        | |     |||
Db     240517 AAAATTCTCAGTATTTTCATAGTAGTTCTTCTATTCGCTGTTGGATGCGGACAGCAAAAA 240576

Qy         85 AATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAA 144
               |      ||   |  |||     ||      |||  | |   |   ||| || ||||
Db     240577 GAGGAGAAAAAAGAAACAAAAGCGGACAATAAAAATCAAGCTATAACAATTAAACACGCT 240636

Qy        145 GAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTT 204
              |||||    |||||  ||  |||||   | |||   || ||||| |||||||    || |
Db     240637 GAAGGGGAAACGAAGTTAGATAAACCAGCAAAAAAAGTAGTTGTACTTGAGTGGGTATAT 240696

Qy        205 GTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAA 264
                ||| |  ||| | ||  | | |||| | || || ||||| || ||   ||| ||  |
Db     240697 TCTGAAGACTTATTAGCACTTGGTGTTCAGCCAGTAGGGATGGCAGACATTAAGAATTAT 240756

Qy        265 AATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGT 324
              |||       | ||  ||  || | |    | |  | |         |||||| ||||||
Db     240757 AATAAATGGGTAAATACAAAAACAAAACCGAGTAAAGATGTTGTAGATGTAGGGACACGT 240816

Qy        325 AAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAAT 384
               | ||||| |||||||||||||| ||     | |||||||| ||||||||| | | |
Db     240817 CAACAACCAAACTTAGAAGAAATTAGCCGTTTAAAACCAGACTTAATTATTACAGCTTCA 240876

Qy        385 AATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGAT---------T 435
                  |    ||||  ||| | || || ||| |  ||||||| || ||  |         |
Db     240877 TTCCGTGGTAAAGCAATTAAAAATGAATTAGAGCAAATTGCACCAACAGTTATGTTTGAT 240936

Qy        436 GAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCA 495
                |   | |||     | |   | | |  |||||  |  ||  | ||||||  |||| ||
Db     240937 CCATCAACAAGCAATAACGATCACTTTGCTGAAATGACAGAAACATTTAAACAAATTGCA 240996

Qy        496 AAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATT 555
              |||||  | || |||||||||||||||||||||   ||||  ||    |||||   | |
Db     240997 AAAGCAGTTGGAAAAGAAGAAGAAGGTAAAAAAGTATTAGCTGATATGGATAAAGCATTC 241056
```

```
Qy        556 GAAGAATATAAAAAAGAAAT 575
              |  ||    |||  | ||||
Db     241057 GCTGATGCAAAAGCAAAAAT 241076


RESULT 30
AE017225_43
WPCOMMENT
  Sequence split into 53 fragments  LOCUS AE017225 Accession AE017225
     Fragment Name      Begin      End
     AE017225_00            1   110000
     AE017225_01       100001   210000
     AE017225_02       200001   310000
     AE017225_03       300001   410000
     AE017225_04       400001   510000
     AE017225_05       500001   610000
     AE017225_06       600001   710000
     AE017225_07       700001   810000
     AE017225_08       800001   910000
     AE017225_09       900001  1010000
     AE017225_10      1000001  1110000
     AE017225_11      1100001  1210000
     AE017225_12      1200001  1310000
     AE017225_13      1300001  1410000
     AE017225_14      1400001  1510000
     AE017225_15      1500001  1610000
     AE017225_16      1600001  1710000
     AE017225_17      1700001  1810000
     AE017225_18      1800001  1910000
     AE017225_19      1900001  2010000
     AE017225_20      2000001  2110000
     AE017225_21      2100001  2210000
     AE017225_22      2200001  2310000
     AE017225_23      2300001  2410000
     AE017225_24      2400001  2510000
     AE017225_25      2500001  2610000
     AE017225_26      2600001  2710000
     AE017225_27      2700001  2810000
     AE017225_28      2800001  2910000
     AE017225_29      2900001  3010000
     AE017225_30      3000001  3110000
     AE017225_31      3100001  3210000
     AE017225_32      3200001  3310000
     AE017225_33      3300001  3410000
     AE017225_34      3400001  3510000
     AE017225_35      3500001  3610000
     AE017225_36      3600001  3710000
     AE017225_37      3700001  3810000
     AE017225_38      3800001  3910000
     AE017225_39      3900001  4010000
     AE017225_40      4000001  4110000
     AE017225_41      4100001  4210000
     AE017225_42      4200001  4310000
     AE017225_43      4300001  4410000
     AE017225_44      4400001  4510000
     AE017225_45      4500001  4610000
     AE017225_46      4600001  4710000
     AE017225_47      4700001  4810000
     AE017225_48      4800001  4910000
     AE017225_49      4900001  5010000
     AE017225_50      5000001  5110000
     AE017225_51      5100001  5210000
     AE017225_52      5200001  5228663
Continuation (44 of 53) of AE017225 from base 4300001 (AE017225 Bacillus anthracis str.

  Query Match             14.2%;  Score 143.2;  DB 1;  Length 110000;
  Best Local Similarity   55.0%;  Pred. No. 5.2e-13;
  Matches  308;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    1;

Qy         25 AAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAAT 84
              |||||| | ||| |  |     || | ||| | ||   ||        | |     |||
Db      33922 AAAATTCTCAGTATTTTCATAGTAGTTTTTCTATTCGCTGTTGGATGCGGACAGCAAAAA 33981
```

```
Qy        85 AATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAA 144
              |      ||   |  |||     ||      |||  | |   |   ||| || ||||
Db     33982 GAGGAGAAAAAAGAAACAAAAGCGGACAATAAAAATCAAGCTATAACAATTAAACACGCT 34041

Qy       145 GAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTT 204
             |||||    |||||  ||  |||||   | |||   || ||||| ||||| |    || |
Db     34042 GAAGGGGAAACGAAGTTAGATAAACCAGCGAAAAAAGTAGTTGTACTTGAATGGGTATAT 34101

Qy       205 GTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAA 264
                || |  ||| | ||  | | |||| | || || ||||| || ||   ||| ||  |
Db     34102 TCAGAAGACTTATTAGCACTTGGTGTTCAGCCAGTAGGGATGGCAGACATTAAGAATTAT 34161

Qy       265 AATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGT 324
             |||       | ||  ||  || | |    | |  | |         ||| || ||||||
Db     34162 AATAAATGGGTAAATACAAAAACAAAACCGAGTAAAGATGTTGTAGATGTCGGGACACGT 34221

Qy       325 AAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAAT 384
              | ||||| |||||||||||||| ||     | |||||||||||||||||  | | |
Db     34222 CAACAACCAAACTTAGAAGAAATTAGCCGTTTAAAACCAGATTTAATTATCACAGCTTCA 34281

Qy       385 AATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGAT---------T 435
                 |    ||||  ||| | || || ||| |  ||||||| || ||  |         |
Db     34282 TTCCGTGGTAAAGCAATTAAAAATGAATTAGAACAAATTGCACCAACAGTTATGTTTGAT 34341

Qy       436 GAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCA 495
               |   | |||     | |   | | |  |||||  |  ||  | ||||||  |||| ||
Db     34342 CCATCAACAAGCAATAACGATCACTTTGCTGAAATGACAGAAACATTTAAACAAATTGCA 34401

Qy       496 AAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATT 555
             |||||  | || |||||||||||||||||||||   ||||  ||    |||||   | |
Db     34402 AAAGCAGTTGGAAAAGAAGAAGAAGGTAAAAAAGTATTAGCTGATATGGATAAAGCATTC 34461

Qy       556 GAAGAATATAAAAAAGAAAT 575
             |  ||    |||  | ||||
Db     34462 GCTGATGCAAAAGCAAAAAT 34481


RESULT 31
AE017334_43
WPCOMMENT
  Sequence split into 53 fragments  LOCUS AE017334 Accession AE017334
     Fragment Name         Begin      End
     AE017334_00               1   110000
     AE017334_01          100001   210000
     AE017334_02          200001   310000
     AE017334_03          300001   410000
     AE017334_04          400001   510000
     AE017334_05          500001   610000
     AE017334_06          600001   710000
     AE017334_07          700001   810000
     AE017334_08          800001   910000
     AE017334_09          900001  1010000
     AE017334_10         1000001  1110000
     AE017334_11         1100001  1210000
     AE017334_12         1200001  1310000
     AE017334_13         1300001  1410000
     AE017334_14         1400001  1510000
     AE017334_15         1500001  1610000
     AE017334_16         1600001  1710000
     AE017334_17         1700001  1810000
     AE017334_18         1800001  1910000
     AE017334_19         1900001  2010000
     AE017334_20         2000001  2110000
     AE017334_21         2100001  2210000
     AE017334_22         2200001  2310000
     AE017334_23         2300001  2410000
     AE017334_24         2400001  2510000
     AE017334_25         2500001  2610000
     AE017334_26         2600001  2710000
     AE017334_27         2700001  2810000
     AE017334_28         2800001  2910000
     AE017334_29         2900001  3010000
```

```
     AE017334_30         3000001  3110000
     AE017334_31         3100001  3210000
     AE017334_32         3200001  3310000
     AE017334_33         3300001  3410000
     AE017334_34         3400001  3510000
     AE017334_35         3500001  3610000
     AE017334_36         3600001  3710000
     AE017334_37         3700001  3810000
     AE017334_38         3800001  3910000
     AE017334_39         3900001  4010000
     AE017334_40         4000001  4110000
     AE017334_41         4100001  4210000
     AE017334_42         4200001  4310000
     AE017334_43         4300001  4410000
     AE017334_44         4400001  4510000
     AE017334_45         4500001  4610000
     AE017334_46         4600001  4710000
     AE017334_47         4700001  4810000
     AE017334_48         4800001  4910000
     AE017334_49         4900001  5010000
     AE017334_50         5000001  5110000
     AE017334_51         5100001  5210000
     AE017334_52         5200001  5227419
Continuation (44 of 53) of AE017334 from base 4300001 (AE017334 Bacillus anthracis str.

  Query Match             14.2%;  Score 143.2;  DB 1;  Length 110000;
  Best Local Similarity   55.0%;  Pred. No. 5.2e-13;
  Matches  308;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    1;

Qy        25 AAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAAT 84
             |||||| | ||| |  |     || | ||| | ||   ||        | |     |||
Db     33360 AAAATTCTCAGTATTTTCATAGTAGTTTTTCTATTCGCTGTTGGATGCGGACAGCAAAAA 33419

Qy        85 AATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAA 144
              |      ||   |  |||     ||      |||  | |   |   ||| || ||||
Db     33420 GAGGAGAAAAAAGAAACAAAAGCGGACAATAAAAATCAAGCTATAACAATTAAACACGCT 33479

Qy       145 GAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTT 204
             |||||    |||||  ||  |||||   | |||   || ||||| ||||| |    || |
Db     33480 GAAGGGGAAACGAAGTTAGATAAACCAGCGAAAAAAGTAGTTGTACTTGAATGGGTATAT 33539

Qy       205 GTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAA 264
                || |  ||| | ||  | | |||| | || || ||||| || ||   ||| ||  |
Db     33540 TCAGAAGACTTATTAGCACTTGGTGTTCAGCCAGTAGGGATGGCAGACATTAAGAATTAT 33599

Qy       265 AATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGT 324
             |||       | ||  ||  || | |    | |  | |         ||| || ||||||
Db     33600 AATAAATGGGTAAATACAAAAACAAAACCGAGTAAAGATGTTGTAGATGTCGGGACACGT 33659

Qy       325 AAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAAT 384
              | ||||| |||||||||||||| ||     | |||||||||||||||||  | | |
Db     33660 CAACAACCAAACTTAGAAGAAATTAGCCGTTTAAAACCAGATTTAATTATCACAGCTTCA 33719

Qy       385 AATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGAT---------T 435
                 |    ||||  ||| | || || ||| |  ||||||| || ||  |         |
Db     33720 TTCCGTGGTAAAGCAATTAAAAATGAATTAGAACAAATTGCACCAACAGTTATGTTTGAT 33779

Qy       436 GAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCA 495
               |   | |||     | |   | | |  |||||  |  ||  | ||||||  |||| ||
Db     33780 CCATCAACAAGCAATAACGATCACTTTGCTGAAATGACAGAAACATTTAAACAAATTGCA 33839

Qy       496 AAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATT 555
             |||||  | || |||||||||||||||||||||   ||||  ||    |||||   | |
Db     33840 AAAGCAGTTGGAAAAGAAGAAGAAGGTAAAAAAGTATTAGCTGATATGGATAAAGCATTC 33899

Qy       556 GAAGAATATAAAAAAGAAAT 575
             |  ||    |||  | ||||
Db     33900 GCTGATGCAAAAGCAAAAAT 33919


RESULT 32
AE017038
LOCUS       AE017038               290029 bp    DNA     linear   BCT 30-APR-2003
```

```
DEFINITION  Bacillus anthracis str. Ames section 15 of 18 of the complete
            genome.
ACCESSION   AE017038 AE016879
VERSION     AE017038.1  GI:30259002
KEYWORDS    .
SOURCE      Bacillus anthracis str. Ames
  ORGANISM  Bacillus anthracis str. Ames
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.
REFERENCE   1  (bases 1 to 290029)
  AUTHORS   Read,T., Peterson,S., Tourasse,N., Baillie,L., Paulsen,I.,
            Nelson,K., Tettelin,H., Fouts,D., Eisen,J., Gill,S., Holtzapple,E.,
            Okstad,O., Helgason,E., Rilstone,J., Wu,M., Kolonay,J., Beanan,M.,
            Dodson,R., Brinkac,L., Gwinn,M., DeBoy,R., Madupu,R., Daugherty,S.,
            Durkin,A., Haft,D., Nelson,W., Peterson,J., Pop,M., Khouri,H.,
            Radune,D., Benton,J., Mahamoud,Y., Jiang,L., Hance,I., Weidman,J.,
            Berry,K., Plaut,R., Wolf,A., Watkins,K., Nierman,W., Hazen,A.,
            Cline,R., Redmond,C., Thwaite,J., White,O., Salzberg,S.,
            Thomason,B., Friedlander,A., Koehler,T., Hanna,P., Kolsto,A.-B. and
            Fraser,C.
  TITLE     The genome sequence of Bacillus anthracis Ames and comparison to
            closely related bacteria
  JOURNAL   Nature 423 (6935), 81-86 (2003)
  MEDLINE   22608414
   PUBMED   12721629
REFERENCE   2  (bases 1 to 290029)
  AUTHORS   Read,T., Peterson,S., Tourasse,N., Baillie,L., Paulsen,I.,
            Nelson,K., Tettelin,H., Fouts,D., Eisen,J., Gill,S., Holtzapple,E.,
            Okstad,O., Helgason,E., Rilstone,J., Wu,M., Kolonay,J., Beanan,M.,
            Dodson,R., Brinkac,L., Gwinn,M., DeBoy,R., Madupu,R., Daugherty,S.,
            Durkin,A., Haft,D., Nelson,W., Peterson,J., Pop,M., Khouri,H.,
            Radune,D., Benton,J., Mahamoud,Y., Jiang,L., Hance,I., Weidman,J.,
            Berry,K., Plaut,R., Wolf,A., Watkins,K., Nierman,W., Hazen,A.,
            Cline,R., Redmond,C., Thwaite,J., White,O., Salzberg,S.,
            Thomason,B., Friedlander,A., Koehler,T., Hanna,P., Kolsto,A.-B. and
            Fraser,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-MAR-2003) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .290029
                     /organism="Bacillus anthracis str. Ames"
                     /mol_type="genomic DNA"
                     /strain="Ames"
                     /db_xref="taxon:198094"
     gene            complement(197. .2317)
                     /locus_tag="BA4497"
     CDS             complement(197. .2317)
                     /locus_tag="BA4497"
                     /note="similar to GB:Y00264, GB:X06989, GB:M24546,
                     GB:M24547, GB:M34862, GB:M34863, GB:M34864, GB:M34865,
                     GB:M34866, GB:M34867, GB:M34868, GB:M34869, GB:M34870,
                     GB:M34871, GB:M34872, GB:M34873, GB:M34874, GB:M34876,
                     GB:M34877, GB:M34878, GB:M34879, GB:M33112, GB:M34875,
                     GB:M15533, GB:M18734, GB:S60721, GB:S61380, GB:S61383,
                     GB:X06981, GB:X06982, GB:M15532, GB:S41242, GB:S41243,
                     GB:S45135, GB:S45136, GB:X13471, GB:X13475, GB:X13466,
                     GB:X13467, GB:X13468, GB:X13469, GB:X13470, GB:X13472,
                     GB:X13473, GB:X13474, GB:X13476, GB:X13477, GB:X13478,
                     GB:X13479, GB:X13487, GB:X13488, GB:X88307, GB:X88557,
                     SP:P05067, PID:178615, PID:516074, PID:609449, PID:871360,
                     PID:929611,  and PID:929612; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="penicillin-binding protein"
                     /protein_id="AAP28208.1"
                     /db_xref="GI:30259003"
                     /translation="MSKKQKQQKKKTHVPFRLNVLFFCVFLMFSAVIVRLGYVQIVRG
                     EEYKNEVERKENSTISNPVPRGKIFDRYGRPVVDNAAVRTITFTKMKGSTAEARLETA
                     KKLADLIEVPTNKLTDRDKKDYWLAIHKEEAKEKITNKDRQEFKDKKIDDKELDERQR
                     NRVTEEEVNQLSAKDLEVLAIKSKMDGGYAMTPQVIKKDATEQEYAIISENLATLPGV
                     DTNVDWDRKYIYDDLFRSVLGGVSTSDEGLPRERLDYYLVRDYNRNDRVGKSYLEQQY
                     EDSLHGNKAEVRNVTDKNGNILETINVSKGQSGNNLNLTIDMELQKRVEEIITKNLLR
                     YKGGQPYLDRAFVVMMNPKNGEVLSMAGKQLVEENGETKVQDFALGTMTSSYPMGSTV
                     KGATVLTGYQTGAIQPGSVQLDEPIRLKGTPPKSSWKTMGYINDLTALKMSSNVYMFK
                     TAMNIAGVQYVRGGTLDIPQKAFDTMRYYFGQFGLGVKTGIDLPNESAGQTGRGNQPG
                     FLLDLSIGQYDTYTPLQLAQYVSTIANGGYRMQPQVVKEIRQPSAKQDEVGKVVHSME
                     PKVLNRIDMPESQIKRVQEGFRQVFNDSGGTAAKYFTGAPYKAAGKTGTAQTVYGGDK
                     EIGKKANGERKETYNLTLVGYAPLEDPEVAFSVVVPWVDDKSGINGYISRDIMDAYFD
                     LKKVENGEATKDEIDKKNKEQDDE"
     gene            complement(2427. .3704)
                     /locus_tag="BA4498"
     CDS             complement(2427. .3704)
                     /locus_tag="BA4498"
                     /note="identified by match to PFAM protein family HMM
                     PF03209"
                     /codon_start=1
                     /transl_table=11
                     /product="membrane protein, putative"
                     /protein_id="AAP28209.1"
                     /db_xref="GI:30259004"
                     /translation="MKWKHIIGDVEVNRDLVLLLLIGGLYTLAISLSNTFVNIYLWKQ
                     TQNYVNLGLYNLASVVLQPLTFLIGGKLAKRIDRSILLRIGVGTLATFFIVVLVAGKS
                     ASHYILLMGALLGVGYGFYWLAFNLLTFEITEPETRDFFNGFLGLLTSFSGMIGPIAA
                     GYTISRMEKWSGYTVIFFLSLVLFAIAVVLSFFLSKRECEGRYEITQVLKERRIDKNW
                     GRITRAHFFQGLREGTFIFVISVYVYLATDSELALGKYSLVNSAVSFVCYYLVARMLK
                     KEWRKKAILLGGIILYAVVFLVIFNVTYTKLLIYAACIAIAYPILLVPYGSMTYDVIG
                     RAKNAREWRVEYVVVRELWLNAGRICSVLSFLCAVLFFPPEKSLPFLLCILGAGHFLI
                     YFAVKNVKYDGGNAGKTSVVAQGTTQNQTEPEG"
     gene            complement(3820. .4431)
                     /gene="sodA-1"
                     /locus_tag="BA4499"
     CDS             complement(3820. .4431)
                     /gene="sodA-1"
                     /locus_tag="BA4499"
                     /note="similar to GB:J05158, SP:P22792, PID:179936,
                     GB:J05392, GB:X60306, GB:Z48199, SP:P18827, PID:338634,
                     PID:36652,  and PID:666052; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="superoxide dismutase, Mn"
                     /protein_id="AAP28210.1"
                     /db_xref="GI:30259005"
                     /translation="MAKHELPNLPYAYDALEPHFDKETMNIHHTKHHNTYITNLNAAL
                     EGHAELADKSVEELVANLNEVPEAIRTAVRNNGGGHANHTFFWTILSPNGGGQPVGEL
                     ATAIEAKFGSFDAFKEEFAKAGATRFGSGWAWLVVNNGELEVTSTPNQDSPLTEGKTP
                     VIGLDVWEHAYYLNYQNRRPDYIGAFWNVVDWNAAEKRYQEAK"
     gene            complement(4653. .4726)
                     /locus_tag="tRNA-Met-5"
     tRNA            complement(4653. .4726)
                     /locus_tag="tRNA-Met-5"
                     /product="tRNA-Met"
                     /note="codon recognized: AUG"
     gene            4969. .5727
                     /locus_tag="BA4500"
     CDS             4969. .5727
                     /locus_tag="BA4500"
                     /note="similar to GB:S67291, GB:S67292, GB:S67294,
                     GB:X51943, GB:M13361, GB:M60515, GB:M60516, GB:M60518,
                     GB:M60519, GB:M60520, GB:M60521, GB:X59065, GB:X65778,
                     GB:X59612, SP:P05230, PID:181942, PID:182559, PID:292022,
                     PID:292027, PID:32436, PID:386768, PID:396164, PID:396166,
                     PID:553170, PID:553171, GB:S67291, GB:S67292, GB:S67294,
                     GB:X51943, GB:M13361, GB:M60515, GB:M60516, GB:M60518,
                     GB:M60519, GB:M60520, GB:M60521, GB:X59065, GB:X65778,
                     GB:X59612, SP:P05230, PID:181942, PID:182559, PID:292022,
                     PID:292027, PID:32436, PID:386768, PID:396164, PID:396166,
                     PID:553170, PID:553171, GB:S67291, GB:S67292, GB:S67294,
                     GB:X51943, GB:M13361, GB:M60515, GB:M60516, GB:M60518,
                     GB:M60519, GB:M60520, GB:M60521, GB:X59065, GB:X65778,
                     GB:X59612, SP:P05230, PID:181942, PID:182559, PID:292022,
                     PID:292027, PID:32436, PID:386768, PID:396164, PID:396166,
                     PID:553170, PID:553171, GB:S67291, GB:S67292, GB:S67294,
                     GB:X51943, GB:M13361, GB:M60515, GB:M60516, GB:M60518,
                      GB:M60519, GB:M60520, GB:M60521, GB:X59065, GB:X65778,
                      GB:X59612, SP:P05"
                      /codon_start=1
                      /transl_table=11
                      /product="conserved hypothetical protein"
                      /protein_id="AAP28211.1"
                      /db_xref="GI:30259006"
                      /translation="MSIFTQLVKSVYSPKDMALFRFQKIGKTILYIMLLCLITTIPRT
                      FFYGSFIQDSVTMVNQAIEKDLPDFKIENGELKADIEEPIQKEEGDALFVFDPNTTDI
                      EKYQNKTGLFILKDKVVSIGNGQTQTYSYNDLLGTSLEKKDLQDFISVFDSIYPILLA
                      VIGVLVYLFQLFITFLGITLLAFIGSAMSGQRKLSYKQVWTLTAYSYTIPTIFFMIMD
                      LFKIVVPGSTFIYIAVVLIVLYLTIKEVPKPNEK"
     gene             5799. .6164
                      /locus_tag="BA4501"
     CDS              5799. .6164
                      /locus_tag="BA4501"
                      /note="identified by Glimmer2; putative"
                      /codon_start=1
                      /transl_table=11
                      /product="conserved hypothetical protein"
                      /protein_id="AAP28212.1"
                      /db_xref="GI:30259007"
                      /translation="MKRLGIFLFVLVLGYIFYYDIKIGTLPMLSSYKKTNAAQTIKQE
                      STNTKQNKENKAEKETDVTYKTIEVKTGETVLSITEQINKKKIPSIEKVIDDFKQLNK
                      STSATKIQIGKSYKFPLYQ"
     gene             6329. .7432
                      /gene="gcpE"
                      /locus_tag="BA4502"
     CDS              6329. .7432
                      /gene="gcpE"
                      /locus_tag="BA4502"
                      /note="identified by match to PFAM protein family HMM
                      PF02784"
                      /codon_start=1
                      /transl_table=11
                      /product="gcpE protein"
                      /protein_id="AAP28213.1"
                      /db_xref="GI:30259008"
                      /translation="MTHRTKTRPVKVGNLTIGGNNELIIQSMTTTKTHDVEATVAEIK
                      RLEEAGCQVVRVAVPDERAANAIADIKKQINIPLVADIHFDYRLALKAIEGGIDKVRI
                      NPGNIGRRHKVEAVVNAAKERGIPIRIGVNAGSLERHILEKYGYPTADGMVESALHHI
                      KILEDLDFHDIIVSMKASDVNLAIEAYEKAARAFDYPLHLGITESGTLFAGTVKSAAG
                      LGAILNKGIGNTLRISLSADPVEEVKVARELLKSFGLASNAATLISCPTCGRIEIDLI
                      SIANEVEEYISTLQVPIKVAVLGCAVNGPGEAREADIGIAGARGEGLLFRKGQVVRKV
                      PEEIMVEELKKEIDVIAAEMAAEREKEKETQEQ"
     gene             complement(7513. .7926)
                      /locus_tag="BA4503"
     CDS              complement(7513. .7926)
                      /locus_tag="BA4503"
                      /note="similar to GB:S67291, GB:S67292, GB:S67294,
                      GB:X51943, GB:M13361, GB:M60515, GB:M60516, GB:M60518,
                      GB:M60519, GB:M60520, GB:M60521, GB:X59065, GB:X65778,
                      GB:X59612, SP:P05230, PID:181942, PID:182559, PID:292022,

  Query Match            14.2%;  Score 143.2;  DB 1;  Length 290029;
  Best Local Similarity  55.0%;  Pred. No. 4.1e-13;
  Matches  308;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    1;

Qy          25 AAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAAT 84
               |||||| | ||| |  |     || | ||| | ||   ||        | |     ||| 
Db      243302 AAAATTCTCAGTATTTTCATAGTAGTTTTTCTATTCGCTGTTGGATGCGGACAGCAAAAA 243361

Qy          85 AATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAA 144
                |      ||   |  |||     ||      |||  | |   |   ||| || ||||  
Db      243362 GAGGAGAAAAAAGAAACAAAAGCGGACAATAAAAATCAAGCTATAACAATTAAACACGCT 243421

Qy         145 GAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTT 204
               |||||    |||||  ||  |||||   | |||   || ||||| ||||| |    || |
Db      243422 GAAGGGGAAACGAAGTTAGATAAACCAGCGAAAAAAGTAGTTGTACTTGAATGGGTATAT 243481

Qy         205 GTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAA 264
                  || |  ||| | ||  | | |||| | || || ||||| || ||   ||| ||  | 
Db      243482 TCAGAAGACTTATTAGCACTTGGTGTTCAGCCAGTAGGGATGGCAGACATTAAGAATTAT 243541

Qy         265 AATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGT 324
               |||       | ||  ||  || | |    | |  | |         ||| || ||||||
Db      243542 AATAAATGGGTAAATACAAAAACAAAACCGAGTAAAGATGTTGTAGATGTCGGGACACGT 243601

Qy         325 AAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAAT 384
                | ||||| |||||||||||||| ||     | |||||||||||||||||  | | |   
Db      243602 CAACAACCAAACTTAGAAGAAATTAGCCGTTTAAAACCAGATTTAATTATCACAGCTTCA 243661

Qy         385 AATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGAT---------T 435
                   |    ||||  ||| | || || ||| |  ||||||| || ||  |         |
Db      243662 TTCCGTGGTAAAGCAATTAAAAATGAATTAGAACAAATTGCACCAACAGTTATGTTTGAT 243721

Qy         436 GAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCA 495
                 |   | |||     | |   | | |  |||||  |  ||  | ||||||  |||| ||
Db      243722 CCATCAACAAGCAATAACGATCACTTTGCTGAAATGACAGAAACATTTAAACAAATTGCA 243781

Qy         496 AAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATT 555
               |||||  | || |||||||||||||||||||||   ||||  ||    |||||   | | 
Db      243782 AAAGCAGTTGGAAAAGAAGAAGAAGGTAAAAAAGTATTAGCTGATATGGATAAAGCATTC 243841

Qy         556 GAAGAATATAAAAAAGAAAT 575
               |  ||    |||  | ||||
Db      243842 GCTGATGCAAAAGCAAAAAT 243861


RESULT 33
AE017355_43
WPCOMMENT
  Sequence split into 53 fragments  LOCUS AE017355 Accession AE017355
     Fragment Name       Begin      End
     AE017355_00             1     110000
     AE017355_01        100001     210000
     AE017355_02        200001     310000
     AE017355_03        300001     410000
     AE017355_04        400001     510000
     AE017355_05        500001     610000
     AE017355_06        600001     710000
     AE017355_07        700001     810000
     AE017355_08        800001     910000
     AE017355_09        900001    1010000
     AE017355_10       1000001    1110000
     AE017355_11       1100001    1210000
     AE017355_12       1200001    1310000
     AE017355_13       1300001    1410000
     AE017355_14       1400001    1510000
     AE017355_15       1500001    1610000
     AE017355_16       1600001    1710000
     AE017355_17       1700001    1810000
     AE017355_18       1800001    1910000
     AE017355_19       1900001    2010000
     AE017355_20       2000001    2110000
     AE017355_21       2100001    2210000
     AE017355_22       2200001    2310000
     AE017355_23       2300001    2410000
     AE017355_24       2400001    2510000
     AE017355_25       2500001    2610000
     AE017355_26       2600001    2710000
     AE017355_27       2700001    2810000
     AE017355_28       2800001    2910000
     AE017355_29       2900001    3010000
     AE017355_30       3000001    3110000
     AE017355_31       3100001    3210000
     AE017355_32       3200001    3310000
     AE017355_33       3300001    3410000
     AE017355_34       3400001    3510000
     AE017355_35       3500001    3610000
     AE017355_36       3600001    3710000
     AE017355_37       3700001    3810000
     AE017355_38       3800001    3910000
     AE017355_39       3900001    4010000
     AE017355_40       4000001    4110000
     AE017355_41       4100001    4210000
```

```
      AE017355_42        4200001    4310000
      AE017355_43        4300001    4410000
      AE017355_44        4400001    4510000
      AE017355_45        4500001    4610000
      AE017355_46        4600001    4710000
      AE017355_47        4700001    4810000
      AE017355_48        4800001    4910000
      AE017355_49        4900001    5010000
      AE017355_50        5000001    5110000
      AE017355_51        5100001    5210000
      AE017355_52        5200001    5237682
Continuation (44 of 53) of AE017355 from base 4300001 (AE017355 Bacillus thuringiensis s

  Query Match             13.9%;  Score 140;  DB 1;  Length 110000;
  Best Local Similarity   54.6%;  Pred. No. 1.6e-12;
  Matches  306;  Conservative    0;  Mismatches  245;  Indels    9;  Gaps    1;

Qy          25 AAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAAT 84
               | |||| | ||| |  |     || |  || | ||   ||        | |     |||
Db       26313 AGAATTCTCAGTATTTTCATAGTAGTTCTTCTATTCGCTGTTGGATGCGGACAGCAAAAA 26372

Qy          85 AATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAA 144
                |      ||   |  |||       ||      |||  | |   |   ||| || ||||
Db       26373 GAGGAGAAAAAAGAAACAAAAGCGGACAATAAAAATCAAGCTATAACAATTAAACACGCT 26432

Qy         145 GAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTT 204
               |||||    |||||  ||  |||||   | |||   || ||||| ||||| |    || |
Db       26433 GAAGGGGAAACGAAGTTAGATAAACCAGCGAAAAAAGTAGTTGTACTTGAATGGGTATAT 26492

Qy         205 GTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAA 264
                  || |  ||| | ||  | | |||| | || || ||||| || ||   ||| ||  |
Db       26493 TCAGAAGACTTATTAGCACTTGGTGTTCAGCCAGTAGGGATGGCAGACATTAAGAATTAT 26552

Qy         265 AATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGT 324
               |||       | ||  ||  || | |    | |  | |         ||| || ||||||
Db       26553 AATAAATGGGTAAATACAAAAACAAAACCGAGTAAAGATGTTGTAGATGTCGGGACACGT 26612

Qy         325 AAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAAT 384
                | ||||| |||||||||||||| ||     | ||||||||||||||||||  | | |
Db       26613 CAACAACCAAACTTAGAAGAAATTAGCCGTTTAAAACCAGATTTAATTATCACAGCTTCA 26672

Qy         385 AATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGAT---------T 435
                   |    ||||  ||| | || || ||| |  ||||||| || ||  |         |
Db       26673 TTCCGTGGTAAAGCAATTAAAAATGAATTAGAACAAATTGCACCAACAGTTATGTTTGAT 26732

Qy         436 GAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCA 495
                 |   | |||     | |   | | |  |||||  |  ||  | ||||||  |||| ||
Db       26733 CCATCAACAAGCAATAACGATCACTTTGCTGAAATGACAGAAACATTTAAACAAATTGCA 26792

Qy         496 AAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATT 555
               |||||  | || |||||||||||||||||||||   ||||  ||    |||||   | |
Db       26793 AAAGCAGTTGGAAAAGAAGAAGAAGGTAAAAAAGTATTAGCTGATATGGATAAAGCATTC 26852

Qy         556 GAAGAATATAAAAAAGAAAT 575
               |  ||    |||  | ||||
Db       26853 GCTGATGCAAAAGCAAAAAT 26872


RESULT 34
BX571857_00/c
WPCOMMENT
  Sequence split into 28 fragments  LOCUS BX571857 Accession BX571857
    Fragment Name       Begin       End
    BX571857_00             1       110000
    BX571857_01        100001       210000
    BX571857_02        200001       310000
    BX571857_03        300001       410000
    BX571857_04        400001       510000
    BX571857_05        500001       610000
    BX571857_06        600001       710000
    BX571857_07        700001       810000
    BX571857_08        800001       910000
    BX571857_09        900001      1010000
```

```
            BX571857_10        1000001    1110000
            BX571857_11        1100001    1210000
            BX571857_12        1200001    1310000
            BX571857_13        1300001    1410000
            BX571857_14        1400001    1510000
            BX571857_15        1500001    1610000
            BX571857_16        1600001    1710000
            BX571857_17        1700001    1810000
            BX571857_18        1800001    1910000
            BX571857_19        1900001    2010000
            BX571857_20        2000001    2110000
            BX571857_21        2100001    2210000
            BX571857_22        2200001    2310000
            BX571857_23        2300001    2410000
            BX571857_24        2400001    2510000
            BX571857_25        2500001    2610000
            BX571857_26        2600001    2710000
            BX571857_27        2700001    2799802
LOCUS       BX571857               2799802 bp    DNA     circular BCT 24-JUN-2004
DEFINITION  Staphylococcus aureus strain MSSA476, complete genome.
ACCESSION   BX571857
VERSION     BX571857.1  GI:49243355
KEYWORDS    complete genome.
SOURCE      Staphylococcus aureus subsp. aureus MSSA476
  ORGANISM  Staphylococcus aureus subsp. aureus MSSA476
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1  (bases 1 to 2799802)
  AUTHORS   Holden,M.T.G., Feil,E.J., Lindsay,J.A., Peacock,S.J., Day,N.P.J.,
            Enright,M.C., Foster,T.J., Moore,C.E., Hurst,L., Atkin,R.,
            Barron,A., Bason,N., Bentley,S.D., Chillingworth,C.,
            Chillingworth,T., Churcher,C., Clark,L., Corton,C., Cronin,A.,
            Doggett,J., Dowd,L., Feltwell,T., Hance,Z., Harris,B., Hauser,H.,
            Holroyd,S., Jagels,K., James,K.D., Lennard,N., Line,A., Mayes,R.,
            Moule,S., Mungall,K., Ormond,D., Quail,M.A., Rabbinowitsch,E.,
            Rutherford,K., Sanders,M., Sharp,S., Simmonds,M., Stevens,K.,
            Whitehead,S., Barrell,B.G., Spratt,B.G. and Parkhill,J.
  TITLE     Complete genomes of two clinical Staphylococcus aureus strains:
            evidence for the rapid evolution of virulence and drug resistance
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (26), 9786-9791 (2004)
REFERENCE   2  (bases 1 to 2799802)
  AUTHORS   Holden,M.T.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-2004) Submitted on behalf of the Pathogen
            Sequencing Unit, Sanger Institute, Wellcome Trust Genome Campus,
            Hinxton, Cambridge CB10 1SA, E-mail: mh3@sanger.ac.uk
FEATURES             Location/Qualifiers
     source          1..2799802
                     /organism="Staphylococcus aureus subsp. aureus MSSA476"
                     /mol_type="genomic DNA"
                     /strain="MSSA476"
                     /db_xref="taxon:282459"
     gene            517..1878
                     /locus_tag="SAS0001"
     CDS             517..1878
                     /locus_tag="SAS0001"
                     /note="Ortholog of S. aureus MRSA252 (BX571856) SAR0001"
                     /codon_start=1
                     /transl_table=11
                     /product="chromosomal replication initiator protein DnaA"
                     /protein_id="CAG41773.1"
                     /db_xref="GI:49243356"
                     /translation="MSEKEIWEKVLEIAQEKLSAVSYSTFLKDTELYTIKDGEAIVLS
                     SIPFNANWLNQQYAEIIQAILFDVVGYEVKPHFITTEELANYSNNETATPKETTKPST
                     ETTEDNHVLGREQFNAHNTFDTFVIGPGNRFPHAASLAVAEAPAKAYNPLFIYGGVGL
                     GKTHLMHAIGHHVLDNNPDAKVIYTSSEKFTNEFIKSIRDNEGEAFRERYRNIDVLLI
                     DDIQFIQNKVQTQEEFFYTFNELHQNNKQIVISSDRPPKEIAQLEDRLRSRFEWGLIV
                     DITPPDYETRMAILQKKIEEEKLDIPPEALNYIANQIQSNIRELEGALTRLLAYSQLL
                     GKPITTELTAEALKDIIQAPKSKKITIQDIQKIVGQYYNVRIEDFSAKKRTKSIAYPR
                     QIAMYLSRELTDFSLPKIGEEFGGRDHTTVIHAHEKISKDLKEDPIFKQEVENLEKEI
                     RNV"
     misc_feature    862..1803
                     /locus_tag="SAS0001"
                     /note="Pfam match to entry PF00308 bac_dnaA, Bacterial
                     dnaA protein, score 681.8, E-value 3.5e-202"
     misc_feature    982. .1005
                     /locus_tag="SAS0001"
                     /note="PS00017 ATP/GTP-binding site motif A (P-loop)."
     misc_feature    1744. .1803
                     /locus_tag="SAS0001"
                     /note="PS01008 DnaA protein signature."
     gene            2156. .3289
                     /locus_tag="SAS0002"
     CDS             2156. .3289
                     /locus_tag="SAS0002"
                     /EC_number="2.7.7.7"
                     /note="Ortholog of S. aureus MRSA252 (BX571856) SAR0002"
                     /codon_start=1
                     /transl_table=11
                     /product="DNA polymerase III, beta chain"
                     /protein_id="CAG41774.1"
                     /db_xref="GI:49243357"
                     /translation="MMEFTIKRDYFITQLNDTLKAISPRTTLPILTGIKIDAKEHEVI
                     LTGSDSEISIEITIPKTVDGEDIVNISETGSVVLPGRFFVDIIKKLPGKDVKLSTNEQ
                     FQTLITSGHSEFNLSGLDPDQYPLLPQVSRDDAIQLSVKVLKNVIAQTNFAVSTSETR
                     PVLTGVNWLIQENELICTATDSHRLAVRKLQLEDVSENKNVIIPGKALAELNKIMSDN
                     EEDIDIFFASNQVLFKVGNVNFISRLLEGHYPDTTRLFPENYEIKLSIDNGEFYHAID
                     RASLLAREGGNNVIKLSTGDDVVELSSTSPEIGTVKEEVDANDVEGGSLKISFNSKYM
                     MDALKAIDNDEVEVEFFGTMKPFILKPKGDDSVTQLILPIRTY"
     misc_feature    2159. .2539
                     /locus_tag="SAS0002"
                     /note="Pfam match to entry PF00712 DNA_pol3_beta, DNA
                     polymerase III beta subunit, N-terminal domain, score
                     224.5, E-value 1.6e-64"
     misc_feature    2564. .2908
                     /locus_tag="SAS0002"
                     /note="Pfam match to entry PF02767 DNA_pol3_beta_2, DNA
                     polymerase III beta subunit, central domain, score 222.4,
                     E-value 7e-64"
     misc_feature    2912. .3280
                     /locus_tag="SAS0002"
                     /note="Pfam match to entry PF02768 DNA_pol3_beta_3, DNA
                     polymerase III beta subunit, C-terminal domain, score
                     209.8, E-value 4.4e-60"
     gene            3670. .3915
                     /locus_tag="SAS0003"
     CDS             3670. .3915
                     /locus_tag="SAS0003"
                     /note="Ortholog of S. aureus MRSA252 (BX571856) SAR0003"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="CAG41775.1"
                     /db_xref="GI:49243358"
                     /translation="MIILVQEVVVEGDINLGQFLKTEGIIESGGQAKWFLQDVEVLIN
                     GVRETRRGKKLEHQDRIDIPELPEDAGSFLIIHQGEQ"
     gene            3912. .5024
                     /locus_tag="SAS0004"
     CDS             3912. .5024
                     /locus_tag="SAS0004"
                     /note="Ortholog of S. aureus MRSA252 (BX571856) SAR0004"
                     /codon_start=1
                     /transl_table=11
                     /product="DNA replication and repair protein RecF"
                     /protein_id="CAG41776.1"
                     /db_xref="GI:49243359"
                     /translation="MKLNTLQLENYRNYDEVTLKCHPDVNILIGENAQGKTNLLESIY
                     TLALAKSHRTSNDKELIRFNADYAKIEGELSYRHGTMPLTMFITKKGKQVKVNHLEQS
                     RLTQYIGHLNVVLFAPEDLNIVKGSPQIRRRFIDMELGQISAVYLNDLAQYQRILKQK
                     NNYLKQLQLGQKKDLTMLEVLNQQFAEYAMKVTDKRAHFIQELESLAKPIHAGITNDK
                     EALSLNYLPSLKFDYAQNEAARLEEIMSILSDNMQREKERGISLFGPHRDDISFDVNG
                     MDAQTYGSQGQQRTTALSIKLAEIELMNIEVGEYPILLLDDVLSELDDSRQTHLLSTI
                     QHKVQTFVTTTSVDGIDHEIMNNAKLYRINQGEIIK"
     misc_feature    3912. .4526
                     /locus_tag="SAS0004"
                     /note="Pfam match to entry PF02463 SMC_N, RecF/RecN/SMC N
                     terminal domain, score -21.4, E-value 2.5e-05"
```

```
     misc_feature    3999. .4022
                     /locus_tag="SAS0004"
                     /note="PS00017 ATP/GTP-binding site motif A (P-loop)."
     misc_feature    4242. .4319
                     /locus_tag="SAS0004"
                     /note="PS00617 RecF protein signature 1."
     misc_feature    4848. .4901
                     /locus_tag="SAS0004"
                     /note="PS00618 RecF protein signature 2."
     gene            5037. .6968
                     /locus_tag="SAS0005"
     CDS             5037. .6968
                     /locus_tag="SAS0005"
                     /EC_number="5.99.1.3"
                     /note="Ortholog of S. aureus MRSA252 (BX571856) SAR0005"
                     /codon_start=1
                     /transl_table=11
                     /product="DNA gyrase subunit B"
                     /protein_id="CAG41777.1"
                     /db_xref="GI:49243360"
                     /translation="MTALSDVNNTDNYGAGQIQVLEGLEAVRKRPGMYIGSTSERGLH
                     HLVWEIVDNSIDEALAGYANQIEVVIEKDNWIKVTDNGRGIPVDIQEKMGRPAVEVIL
                     TVLHAGGKFGGGGYKVSGGLHGVGSSVVNALSQDLEVYVHRNETIYHQAYKKGVPQFD
                     LKEVGTTDKTGTVIRFKADGEIFTETTVYNYETLQQRIRELAFLNKGIQITLRDERDE
                     ENVREDSYHYEGGIKSYVELLNENKEPIHDEPIYIHQSKDDIEVEIAIQYNSGYATNL
                     LTYANNIHTYEGGTHEDGFKRALTRVLNSYGLSSKIMKEEKDRLSGEDTREGMTAIIS
                     IKHGDPQFEGQTKTKLGNSEVRQVVDKLFSEHFERFLYENPQVARTVVEKGIMAARAR
                     VAAKKAREVTRRKSALDVASLPGKLADCSSKSPEECEIFLVEGDSAGGSTKSGRDSRT
                     QAILPLRGKILNVEKARLDRILNNNEIRQMITAFGTGIGGDFDLAKARYHKIVIMTDA
                     DVDGAHIRTLLLTFFYRFMRPLIEAGYVYIAQPPLYKLTQGKQKYYVYNDRELDKLKS
                     ELNPTPKWSIARYKGLGEMNADQLWETTMNPEHRALLQVKLEDAIEADQTFEMLMGDV
                     VENRRQFIEDNAVYANLDF"
     misc_feature    5148. .5579
                     /locus_tag="SAS0005"
                     /note="Pfam match to entry PF02518 HATPase_c, Histidine
                     kinase-, DNA gyrase B-, and HSP90-like ATPase, score
                     103.9, E-value 3.4e-28"
     misc_feature    5538. .5558
                     /locus_tag="SAS0005"
                     /note="PS00154 E1-E2 ATPases phosphorylation site."
     misc_feature    5727. .6239

  Query Match             12.5%;  Score 125.6;  DB 1;  Length 110000;
  Best Local Similarity   50.9%;  Pred. No. 2.4e-10;
  Matches  413;  Conservative    0;  Mismatches  379;  Indels   20;  Gaps    4;

Qy          27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
               ||||  ||   |  |||      |   | |  || |||     |||| |||| |    ||
Db      104242 AATTAAAATGCTTGTTGTTACGCTTGCTTTCCTACTTGTTTTAGCAGGATGTAGTGGGAA 104183

Qy          87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
               |      |   | |||  | |  |   |   || ||   |  |  ||| || || | |
Db      104182 TTCAAATAAACAATCATCTGATAACAAAGATAAGGAAACAACTTCAATTAAACATGCAAT 104123

Qy         147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
                ||||| ||  || |           | || || |||||||||   ||  |  |
Db      104122 GGGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCAC 104063

Qy         207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
               ||| |     ||  ||||||| ||||||||||||| ||    |  ||         | | ||||
Db      104062 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAA 104003

Qy         267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
                |  |   |  ||    ||| | || |  |   | |    |    ||||||   |
Db      104002 ACCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACC 103943

Qy         327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                  |||||||||||| ||||||  |||| | ||||| || |||||| | ||    ||
Db      103942 TGCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGT 103883

Qy         387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
               |||| |  ||    ||||  |  | |||  |||||| || ||             || ||
Db      103882 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAG 103836
```

```
Qy        447 TTTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAG 505
              |||| |  || |   |  | ||| | |  ||| |   |||   |||    ||||||||||
Db     103835 TTTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAG 103776

Qy        506 GTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATA 565
              | |||||| |||||| | || |     |  || |  ||||| | ||| | |  | ||
Db     103775 GGAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCC 103716

Qy        566 AAAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAAT 622
              |||||||   || | | | | ||||| ||  |  |  ||||   ||   || || ||| |
Db     103715 AAAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACT 103656

Qy        623 CAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
                 ||     || |||| |||| |  | |  || |||  |              ||  |
Db     103655 TCCGT---GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 103599

Qy        683 TTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTT 742
               |  || |     | || || | |      ||     |  | |  ||  |||||    |
Db     103598 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 103539

Qy        743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGA 802
               |  ||  |    ||  ||||   |  |   | ||||  |||| | ||| ||  || | |
Db     103538 TCCAACTTACATCCAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAA 103479

Qy        803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
               |    |  ||| | | || ||  || || ||
Db     103478 AATCAGATCCAAATGCGAAAGATGCTGCATTA 103447


RESULT 35
BX571857_01/c
WPCOMMENT
  Sequence split into 28 fragments  LOCUS BX571857 Accession BX571857
    Fragment Name        Begin        End
       BX571857_00           1     110000
       BX571857_01      100001     210000
       BX571857_02      200001     310000
       BX571857_03      300001     410000
       BX571857_04      400001     510000
       BX571857_05      500001     610000
       BX571857_06      600001     710000
       BX571857_07      700001     810000
       BX571857_08      800001     910000
       BX571857_09      900001    1010000
       BX571857_10     1000001    1110000
       BX571857_11     1100001    1210000
       BX571857_12     1200001    1310000
       BX571857_13     1300001    1410000
       BX571857_14     1400001    1510000
       BX571857_15     1500001    1610000
       BX571857_16     1600001    1710000
       BX571857_17     1700001    1810000
       BX571857_18     1800001    1910000
       BX571857_19     1900001    2010000
       BX571857_20     2000001    2110000
       BX571857_21     2100001    2210000
       BX571857_22     2200001    2310000
       BX571857_23     2300001    2410000
       BX571857_24     2400001    2510000
       BX571857_25     2500001    2610000
       BX571857_26     2600001    2710000
       BX571857_27     2700001    2799802
 Continuation (2 of 28) of BX571857 from base 100001 (BX571857 Staphylococcus aureus stra

   Query Match             12.5%;  Score 125.6;  DB 1;  Length 110000;
   Best Local Similarity   50.9%;  Pred. No. 2.4e-10;
   Matches  413;  Conservative    0;  Mismatches  379;  Indels   20;  Gaps    4;

Qy         27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
              ||||  ||   |  |||      |   | |  || |||     |||| |||| |    ||
Db       4242 AATTAAAATGCTTGTTGTTACGCTTGCTTTCCTACTTGTTTTAGCAGGATGTAGTGGGAA 4183

Qy         87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
              |      |   | |||  | |  |   |   || ||   |  |  ||| || || | |
Db       4182 TTCAAATAAACAATCATCTGATAACAAAGATAAGGAAACAACTTCAATTAAACATGCAAT 4123

Qy        147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
               ||||| ||  || |        | || || |||||||||   ||  |  |
Db       4122 GGGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCAC 4063

Qy        207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
              ||| |     ||  |||||| |||||||||||| ||    |  ||       | | ||||
Db       4062 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAA 4003

Qy        267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
               |  |   |  ||    ||| | || |  |   | |    |    ||||||   |
Db       4002 ACCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACC 3943

Qy        327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                 |||||||||||| ||||||  |||| | ||||| || |||||| | ||    ||
Db       3942 TGCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGT 3883

Qy        387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
              |||| |  ||    ||||  |  | |||  |||||| || ||             || ||
Db       3882 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAG 3836

Qy        447 TTTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAG 505
              |||| |  || |   |  | ||| | |  ||| |   |||   |||    ||||||||||
Db       3835 TTTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAG 3776

Qy        506 GTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATA 565
              | |||||| |||||| | || |     |  || |  ||||| | ||| | |  | ||
Db       3775 GGAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCC 3716

Qy        566 AAAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAAT 622
              |||||||   || | | | | ||||| ||  |  |  ||||   ||   || || ||| |
Db       3715 AAAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACT 3656

Qy        623 CAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
                 ||     || |||| |||| |  | |  || |||  |              ||  |
Db       3655 TCCGT---GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 3599

Qy        683 TTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTT 742
               |  || |     | || || | |      ||     |  | |  ||  |||||    |
Db       3598 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 3539

Qy        743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGA 802
               |  ||  |    ||  ||||   |  |   | ||||  |||| | ||| ||  || | |
Db       3538 TCCAACTTACATCCAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAA 3479

Qy        803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
               |    |  ||| | | || ||  || || ||
Db       3478 AATCAGATCCAAATGCGAAAGATGCTGCATTA 3447


RESULT 36
AP004822/c
LOCUS       AP004822               290250 bp    DNA     linear   BCT 20-DEC-2002
DEFINITION  Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,
            strain:MW2, section 1/10.
ACCESSION   AP004822 BA000033
VERSION     AP004822.1  GI:21203164
KEYWORDS    .
SOURCE      Staphylococcus aureus subsp. aureus MW2
  ORGANISM  Staphylococcus aureus subsp. aureus MW2
            Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE   1
  AUTHORS   Baba,T., Takeuchi,F., Kuroda,M., Yuzawa,H., Aoki,K., Oguchi,A.,
            Nagai,Y., Iwama,N., Asano,K., Naimi,T., Kuroda,H., Cui,L.,
            Yamamoto,K. and Hiramatsu,K.
  TITLE     Genome and virulence determinants of high virulence
            community-acquired MRSA
  JOURNAL   Lancet 359 (9320), 1819-1827 (2002)
  MEDLINE   22040717
   PUBMED   12044378
```

```
REFERENCE   2  (bases 1 to 290250)
  AUTHORS   Director-General, Biotechnology Center, Aoki,K., Oguchi,A.,
            Nagai,Y., Asano,K., Iwama,N., Baba,T., Kuroda,M., Hiramatsu,K. and
            Kikuchi,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAR-2002) Director-General, Biotechnology Center,
            National Institute of Technology and Evaluation, Biotechnology
            Center; 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
            (E-mail:bio@nite.go.jp, URL:http://www.bio.nite.go.jp/,
            Tel:81-3-3481-1933, Fax:81-3-3481-8424)
FEATURES             Location/Qualifiers
     source          1. .290250
                     /organism="Staphylococcus aureus subsp. aureus MW2"
                     /mol_type="genomic DNA"
                     /strain="MW2"
                     /db_xref="taxon:196620"
     gene            517. .1878
                     /gene="dnaA"
     CDS             517. .1878
                     /gene="dnaA"
                     /note="ORFID:MW0001"
                     /codon_start=1
                     /transl_table=11
                     /product="chromosomal replication initiator protein"
                     /protein_id="BAB93866.1"
                     /db_xref="GI:21203165"
                     /translation="MSEKEIWEKVLEIAQEKLSAVSYSTFLKDTELYTIKDGEAIVLS
                     SIPFNANWLNQQYAEIIQAILFDVVGYEVKPHFITTEELANYSNNETATPKETTKPST
                     ETTEDNHVLGREQFNAHNTFDTFVIGPGNRFPHAASLAVAEAPAKAYNPLFIYGGVGL
                     GKTHLMHAIGHHVLDNNPDAKVIYTSSEKFTNEFIKSIRDNEGEAFRERYRNIDVLLI
                     DDIQFIQNKVQTQEEFFYTFNELHQNNKQIVISSDRPPKEIAQLEDRLRSRFEWGLIV
                     DITPPDYETRMAILQKKIEEEKLDIPPEALNYIANQIQSNIRELEGALTRLLAYSQLL
                     GKPITTELTAEALKDIIQAPKSKKITIQDIQKIVGQYYNVRIEDFSAKKRTKSIAYPR
                     QIAMYLSRELTDFSLPKIGEEFGGRDHTTVIHAHEKISKDLKEDPIFKQEVENLEKEI
                     RNV"
     gene            2156. .3289
                     /gene="dnaN"
     CDS             2156. .3289
                     /gene="dnaN"
                     /note="ORFID:MW0002"
                     /codon_start=1
                     /transl_table=11
                     /product="DNA polymerase III beta chain"
                     /protein_id="BAB93867.1"
                     /db_xref="GI:21203166"
                     /translation="MMEFTIKRDYFITQLNDTLKAISPRTTLPILTGIKIDAKEHEVI
                     LTGSDSEISIEITIPKTVDGEDIVNISETGSVVLPGRFFVDIIKKLPGKDVKLSTNEQ
                     FQTLITSGHSEFNLSGLDPDQYPLLPQVSRDDAIQLSVKVLKNVIAQTNFAVSTSETR
                     PVLTGVNWLIQENELICTATDSHRLAVRKLQLEDVSENKNVIIPGKALAELNKIMSDN
                     EEDIDIFFASNQVLFKVGNVNFISRLLEGHYPDTTRLFPENYEIKLSIDNGEFYHAID
                     RASLLAREGGNNVIKLSTGDDVVELSSTSPEIGTVKEEVDANDVEGGSLKISFNSKYM
                     MDALKAIDNDEVEVEFFGTMKPFILKPKGDDSVTQLILPIRTY"
     gene            3670. .3915
                     /gene="MW0003"
     CDS             3670. .3915
                     /gene="MW0003"
                     /note="ORFID:MW0003"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB93868.1"
                     /db_xref="GI:21203167"
                     /translation="MIILVQEVVVEGDINLGQFLKTEGIIESGGQAKWFLQDVEVLIN
                     GVRETRRGKKLEHQDRIDIPELPEDAGSFLIIHQGEQ"
     gene            3912. .5024
                     /gene="recF"
     CDS             3912. .5024
                     /gene="recF"
                     /note="ORFID:MW0004"
                     /codon_start=1
                     /transl_table=11
                     /product="DNA repair and genetic recombination protein"
                     /protein_id="BAB93869.1"
                     /db_xref="GI:21203168"
                     /translation="MKLNTLQLENYRNYDEVTLKCHPDVNILIGENAQGKTNLLESIY
                     TLALAKSHRTSNDKELIRFNADYAKIEGELSYRHGTMPLTMFITKKGKQVKVNHLEQS
                     RLTQYIGHLNVVLFAPEDLNIVKGSPQIRRRFIDMELGQISAVYLNDLAQYQRILKQK
                     NNYLKQLQLGQKKDLTMLEVLNQQFAEYAMKVTDKRAHFIQELESLAKPIHAGITNDK
                     EALSLNYLPSLKFDYAQNEAARLEEIMSILSDNMQREKERGISLFGPHRDDISFDVNG
                     MDAQTYGSQGQQRTTALSIKLAEIELMNIEVGEYPILLLDDVLSELDDSRQTHLLSTI
                     QHKVQTFVTTTSVDGIDHEIMNNAKLYRINQGEIIK"
     gene            5034. .6968
                     /gene="gyrB"
     CDS             5034. .6968
                     /gene="gyrB"
                     /note="ORFID:MW0005"
                     /codon_start=1
                     /transl_table=11
                     /product="DNA gyrase subunit B"
                     /protein_id="BAB93870.1"
                     /db_xref="GI:21203169"
                     /translation="MVTALSDVNNTDNYGAGQIQVLEGLEAVRKRPGMYIGSTSERGL
                     HHLVWEIVDNSIDEALAGYANQIEVVIEKDNWIKVTDNGRGIPVDIQEKMGRPAVEVI
                     LTVLHAGGKFGGGGYKVSGGLHGVGSSVVNALSQDLEVYVHRNETIYHQAYKKGVPQF
                     DLKEVGTTDKTGTVIRFKADGEIFTETTVYNYETLQQRIRELAFLNKGIQITLRDERD
                     EENVREDSYHYEGGIKSYVELLNENKEPIHDEPIYIHQSKDDIEVEIAIQYNSGYATN
                     LLTYANNIHTYEGGTHEDGFKRALTRVLNSYGLSSKIMKEEKDRLSGEDTREGMTAII
                     SIKHGDPQFEGQTKTKLGNSEVRQVVDKLFSEHFERFLYENPQVARTVVEKGIMAARA
                     RVAAKKAREVTRRKSALDVASLPGKLADCSSKSPEECEIFLVEGDSAGGSTKSGRDSR
                     TQAILPLRGKILNVEKARLDRILNNNEIRQMITAFGTGIGGDFDLAKARYHKIVIMTD
                     ADVDGAHIRTLLLTFFYRFMRPLIEAGYVYIAQPPLYKLTQGKQKYYVYNDRELDKLK
                     SELNPTPKWSIARYKGLGEMNADQLWETTMNPEHRALLQVKLEDAIEADQTFEMLMGD
                     VVENRRQFIEDNAVYANLDF"
     gene            7005. .9668
                     /gene="gyrA"
     CDS             7005. .9668
                     /gene="gyrA"
                     /note="ORFID:MW0006"
                     /codon_start=1
                     /transl_table=11
                     /product="DNA gyrase subunit A"
                     /protein_id="BAB93871.1"
                     /db_xref="GI:21203170"
                     /translation="MAELPQSRINERNITSEMRESFLDYAMSVIVARALPDVRDGLKP
                     VHRRILYGLNEQGMTPDKSYKKSARIVGDVMGKYHPHGDSSIYEAMVRMAQDFSYRYP
                     LVDGQGNFGSMDGDGAAAMRYTEARMTKITLELLRDINKDTIDFIDNYDGNEREPSVL
                     PARFPNLLANGASGIAVGMATNIPPHNLTELINGVLSLSKNPDISIAELMEDIEGPDF
                     PTAGLILGKSGIRRAYETGRGSIQMRSRAVIEERGGGRQRIVVTEIPFQVNKARMIEK
                     IAELVRDKKIDGITDLRDETSLRTGVRVVIDVRKDANASVILNNLYKQTPLQTSFGVN
                     MIALVNGRPKLINLKEALVHYLEHQKTVVRRRTQYNLRKAKDRAHILEGLRIALDHID
                     EIISTIRESDTDKVAMESLQQRFKLSEKQAQAILDMRLRRLTGLERDKIEAEYNELLN
                     YISELEAILADEEVLLQLVRDELTEIRDRFGDDRRTEIQLGGFEDLEDEDLIPEEQIV
                     ITLSHNNYIKRLPVSTYRAQNRGGRGVQGMNTLEEDFVSQLVTLSTHDHVLFFTNKGR
                     VYKLKGYEVPELSRQSKGIPVVNAIELENDEVISTMIAVKDLESEDNFLVFATKRGVV
                     KRSALSNFSRINRNGKIAISFREDDELIAVRLTSGQEDILIGTSHASLIRFPESTLRP
                     LGRTATGVKGITLREGDEVVGLDVAHANSVDEVLVVTENGYGKRTPVNDYRLSNRGGK
                     GIKTATITERNGNVVCITTVTGEEDLMIVTNAGVIIRLDVADISQNGRAAQGVRLIRL
                     GDDQFVSTVAKVKEDAEDETNEDEQSTSTVSEDGTEQQREAVVNDETPGNAIHTEVID
                     SEENDEDGRIEVRQDFMDRVEEDIQQSSDEE"
     gene            complement(9755. .10567)
                     /gene="MW0007"
     CDS             complement(9755. .10567)
                     /gene="MW0007"
                     /note="ORFID:MW0007"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="BAB93872.1"
                     /db_xref="GI:21203171"
                     /translation="METLNSINIPKRKEDSHKGDYGKILLIGGSANLGGAIMLAARAC
                     VFSGSGLITVATHPTNHSALHSRCPEAMVIDINDTKMLTKMIEMTDSILIGPGLGVDF
                     KGNNAITFLLQNIQPHQNLIVDGDAITIFSKLKPQLPTCRVIFTPHLKEWERLSGIPI
                     EEQTYERNREAVDRLGATVVLKKHGTEIFFKDEDFKLTIGSPAMATGGMGDTLAGMIT
                     SFVGQFDNLKEAVMSATYTHSFIGENLAKDMYVVPPSRLINEIPYAMKQLES"
     gene            10893. .12407
                     /gene="hutH"
     CDS             10893. .12407
                     /gene="hutH"
                     /note="ORFID:MW0008"
                     /codon_start=1
                     /transl_table=11
                     /product="histidine ammonia-lyase"
                     /protein_id="BAB93873.1"
                     /db_xref="GI:21203172"
                     /translation="MTLYLDGETLTIEDIKSFLQQQSKIEIIDDALERVKKSRAVVER
                     IIENEETVYGITTGFGLFSDVRIDPTQYNELQVNLIRSHACGLGEPFSKEVALVMMIL
                     RLNTLLKGHSGATLELVRQLQFFINERIIPIIPQQGSLGASGDLAPLSHLALALIGEG
                     KVLYRGEEKDSDDVLRELNRQPLNLQAKEGLALINGTQAMTAQGVISYIESEDLGYQS
                     EWIAALTHQSLNGIIDAYRHDVHAVRNFQEQINVAARMRDWLEGSTLTTRQAEIRVQD
                     AYTLRCIPQIHGASFQVFNYVKQQLEFEMNAANDNPLIFEEANETFVISGGNFHGQPI
                     AFALDHLKLGVSELANVSERRLERLVNPQLNGDLPAFLSPEPGLQSGAMIMQYAAASL
                     VSENKTLAHPASVDSITSSANQEDHVSMGTTAARHGYQIIENARRVLAIECVIALQAA
                     ELKGVEGLSPKTRRKYEEFRSIVPSITHDRQFHKDIEAVAQYLKQSIYQTTACH"
     gene            12786. .14072
                     /gene="serS"
     CDS             12786. .14072
                     /gene="serS"
                     /note="ORFID:MW0009"
                     /codon_start=1
                     /transl_table=11
                     /product="seryl-tRNA synthetase"
                     /protein_id="BAB93874.1"
                     /db_xref="GI:21203173"
                     /translation="MLDIRLFRNEPVTVKSKIELRGDDPKVVDEILELDEQRRKLISA
                     TEEMKARRNKVSEEIALKKRNKENADDVIAEMRTLGDDIKEKDSQLNEIDNKMTGILC
                     RIPNLISDDVPQGESDEDNVEVKKWGTPREFSFEPKAHWDIVEELKMADFDRAAKVSG
                     ARFVYLTNEGAQLERALMNYMITKHTTQHGYTEMMVPQLVNADTMYGTGQLPKFEEDL
                     FKVEKEGLYTIPTAEVPLTNFYRNEIIQPGVLPEKFTGQSACFRSEAGSAGRDTRGLI
                     RLHQFDKVEMVRFEQPEDSWNALEEMTTNAEAILEELGLPYRRVILCTGDIGFSASKT
                     YDLEVWLPSYNDYKEISSCSNCTDFQARRANIRFKRDKAAKPELAHTLNGSGLAVGRT
                     FAAIVENYQNEDGTVTIPEALVPFMGGKTQISKPVK"
     gene            14724. .15419

  Query Match             12.5%;  Score 125.6;  DB 1;  Length 290250;
  Best Local Similarity   50.9%;  Pred. No. 1.9e-10;
  Matches  413;  Conservative    0;  Mismatches  379;  Indels   20;  Gaps    4;

Qy           27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
                ||||  ||   |  |||      |   | |  || |||     |||| |||| |    ||
Db       105536 AATTAAAATGCTTGTTGTTACGCTTGCTTTCCTACTTGTTTTAGCAGGATGTAGTGGGAA 105477

Qy           87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
                |      |   | |||  | |  |   |   || ||   |  |  ||| || || | |
Db       105476 TTCAAATAAACAATCATCTGATAACAAAGATAAGGAAACAACTTCAATTAAACATGCAAT 105417

Qy          147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
                 ||||| ||  || |        | || || |||||||||   ||  |  |
Db       105416 GGGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCAC 105357

Qy          207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
                ||| |     ||  |||||| ||||||||||||| ||    |  ||         | | ||||
Db       105356 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAA 105297

Qy          267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
                 |  |   |  ||    ||| | || |  |   | |    |    ||||||   |
Db       105296 ACCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACC 105237

Qy          327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                   |||||||||||| ||||||  |||| | ||||| || |||||| | | ||    ||
Db       105236 TGCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGT 105177

Qy          387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
                |||| |  ||    ||||  |  | |||  |||||| || ||             || ||
Db       105176 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAG 105130

Qy          447 TTTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAG 505
                |||| |  || |   |  | ||| | |  ||| |   |||   |||    ||||||||||
Db       105129 TTTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAG 105070
```

```
Qy          506 GTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATA 565
                | |||||| |||||| | || |     |  || |  ||||| | ||| | |  | ||
Db       105069 GGAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCC 105010

Qy          566 AAAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAAT 622
                |||||||   || | | | | ||||| ||  |  |  |||   ||   || || ||| |
Db       105009 AAAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACT 104950

Qy          623 CAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
                   ||     || |||| |||| |  | |  || |||  |              ||  |
Db       104949 TCCGT---GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 104893

Qy          683 TTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTT 742
                 |  || |     | || || | |        ||     |  | |  ||  |||||    |
Db       104892 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 104833

Qy          743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGA 802
                 |  ||  |    ||  ||||   |  |   | ||||  ||||| | ||| ||  || | |
Db       104832 TCCAACTTACATCCAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAA 104773

Qy          803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
                 |    |  ||| | | || ||  || || ||
Db       104772 AATCAGATCCAAATGCGAAAGATGCTGCATTA 104741


RESULT 37
AR354486
LOCUS       AR354486                2115 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 604 from patent US 6593114.
ACCESSION   AR354486
VERSION     AR354486.1  GI:33760570
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2115)
  AUTHORS   Kunsch,C.A., Choi,G.H., Barash,S., Dillon,P.J., Fannon,M.R. and
            Rosen,C.A.
  TITLE     Staphylococcus aureus polynucleotides and sequences
  JOURNAL   Patent: US 6593114-A 604 15-JUL-2003;
FEATURES             Location/Qualifiers
     source          1. .2115
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             12.4%;  Score 125.4;  DB 6;  Length 2115;
  Best Local Similarity   53.2%;  Pred. No. 7.2e-10;
  Matches  289;  Conservative    1;  Mismatches  247;  Indels    6;  Gaps    1;

Qy          458 ATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAG 517
                ||||||||| ||||||||| || ||||||||||| |  | ||  |||  |||| |||||  |||
Db            1 ATTATAATGCAAATATTGAAGCATTTAAAACAGTCGCTAAAGCAGTAGGCAAAGAGAAAG 60

Qy          518 AAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAA 577
                ||||  | || ||  | ||| | || |||| | | | | |  ||   || |||  ||||
Db           61 AAGGCGAGAAGCGTCTGGAAAAGCATGATAAAATATTAGCGGAGATTAGAAAGAAAATTG 120

Qy          578 CTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTC 637
                    |  ||      ||   | |||  |    |      | | | ||||| || ||  |
Db          121 AACAGAGTACGTTAAAATCTGCATTTGCATTCGGTATCTCAAGAGCAGGTATGTTTATTA 180

Qy          638 ATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAA 697
                ||       |  | | | | || |||||| ||| | || | ||| || ||   | | | |
Db          181 ATAATGAAGATACATTTATGGGACAATTCTTAATTAAAATGGGTATTCAACCTGAAGTCA 240

Qy          698 GTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACA 757
                   | :|    ||||    |  || |  |          || |||||  ||  |  | || |
Db          241 MAAAARAMAAAACTACGCATGTTGGTGAACGCAAGGGTGGTCCTTATATATATTTAAATA 300

Qy          758 CTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTT 817
                 ||||    |  |  |  | |||||  |   |||| |  |  ||   ||  ||  |  ||
Db          301 ATGAAGAACTTGCCAATATCAATCCAAAAGTTATGATTTTAGCCACTGACGGAAAAACGG 360
```

```
Qy        818 CTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTG 877
                ||  |         ||   || |     |   | | || |||||   |||||| ||||
Db        361 ACAAAAA------TAGAACGAAATTCATTGATCCTGCAGTTTGGAAATCATTAAAAGCTG 414

Qy        878 TGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAA 937
              ||||| || |    ||| ||    | |||||  |    |||   | |||| | ||  | |
Db        415 TGAAAGATAACAAAGTTTATGACGTTGACCGAAATAAGTGGTTGAAATCAAGGGGGATTA 474

Qy        938 TTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAG 997
              |  |    |||   |||||| ||||  | |   || |  |  | ||||  |   |||||
Db        475 TCGCAAGTGAAAGTATGGCAGAAGATTTAGAAAAAATTGCAGAAAAAGCAAAATAAAAAT 534

Qy        998 ATA 1000
              | |
Db        535 ACA 537

RESULT 38
AR536042
LOCUS       AR536042                2115 bp    DNA     linear   PAT 08-OCT-2004
DEFINITION  Sequence 604 from patent US 6737248.
ACCESSION   AR536042
VERSION     AR536042.1  GI:53927259
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2115)
  AUTHORS   Kunsch,C.A., Choi,G.A., Barash,S.C., Dillon,P.J., Fannon,M.R. and
            Rosen,C.A.
  TITLE     Staphylococcus aureus polynucleotides and sequences
  JOURNAL   Patent: US 6737248-A 604 18-MAY-2004;
FEATURES             Location/Qualifiers
     source          1..2115
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             12.4%;  Score 125.4;  DB 6;  Length 2115;
  Best Local Similarity   53.2%;  Pred. No. 7.2e-10;
  Matches  289;  Conservative    1;  Mismatches  247;  Indels    6;  Gaps    1;

Qy        458 ATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAG 517
              ||||||||| ||||||||| || ||||||||| |  | |||||  |||| |||||  |||
Db          1 ATTATAATGCAAATATTGAAGCATTTAAAACAGTCGCTAAAGCAGTAGGCAAAGAGAAAG 60

Qy        518 AAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAA 577
              ||||  | || ||  | ||| | || ||||| | | | |  ||   || |||  ||||
Db         61 AAGGCGAGAAGCGTCTGGAAAAGCATGATAAAATATTAGCGGAGATTAGAAAGAAAATTG 120

Qy        578 CTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTC 637
                  |  ||      ||   | |||  |    |      | | | ||||| || ||  |
Db        121 AACAGAGTACGTTAAAATCTGCATTTGCATTCGGTATCTCAAGAGCAGGTATGTTTATTA 180

Qy        638 ATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAA 697
              ||       |  | | | | || |||||| ||| | || | ||| || ||   | | | |
Db        181 ATAATGAAGATACATTTATGGGACAATTCTTAATTAAAATGGGTATTCAACCTGAAGTCA 240

Qy        698 GTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACA 757
                 | :|    ||||    | |  || |          || |||||  || |  | || |
Db        241 MAAAARAMAAAACTACGCATGTTGGTGAACGCAAGGGTGGTCCTTATATATATTTAAATA 300

Qy        758 CTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTT 817
               ||||    |  |  |  | |||||  |   |||| |  ||   ||  ||  |  ||
Db        301 ATGAAGAACTTGCCAATATCAATCCAAAAGTTATGATTTTAGCCACTGACGGAAAAACGG 360

Qy        818 CTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTG 877
                ||  |         ||   || |     |   | | || |||||   |||||| ||||
Db        361 ACAAAAA------TAGAACGAAATTCATTGATCCTGCAGTTTGGAAATCATTAAAAGCTG 414

Qy        878 TGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAA 937
              ||||| || |    ||| ||    | |||||  |    |||   | |||| | ||  | |
Db        415 TGAAAGATAACAAAGTTTATGACGTTGACCGAAATAAGTGGTTGAAATCAAGGGGGATTA 474

Qy        938 TTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAG 997
              |  |    |||   |||||| ||||  | |   || |  |  | ||||  |   |||||
Db        475 TCGCAAGTGAAAGTATGGCAGAAGATTTAGAAAAAATTGCAGAAAAAGCAAAATAAAAAT 534

Qy        998 ATA 1000
              | |
Db        535 ACA 537

RESULT 39
AE017355_06/c
WPCOMMENT
  Sequence split into 53 fragments  LOCUS AE017355 Accession AE017355
     Fragment Name      Begin     End
     AE017355_00            1     110000
     AE017355_01       100001     210000
     AE017355_02       200001     310000
     AE017355_03       300001     410000
     AE017355_04       400001     510000
     AE017355_05       500001     610000
     AE017355_06       600001     710000
     AE017355_07       700001     810000
     AE017355_08       800001     910000
     AE017355_09       900001    1010000
     AE017355_10      1000001    1110000
     AE017355_11      1100001    1210000
     AE017355_12      1200001    1310000
     AE017355_13      1300001    1410000
     AE017355_14      1400001    1510000
     AE017355_15      1500001    1610000
     AE017355_16      1600001    1710000
     AE017355_17      1700001    1810000
     AE017355_18      1800001    1910000
     AE017355_19      1900001    2010000
     AE017355_20      2000001    2110000
     AE017355_21      2100001    2210000
     AE017355_22      2200001    2310000
     AE017355_23      2300001    2410000
     AE017355_24      2400001    2510000
     AE017355_25      2500001    2610000
     AE017355_26      2600001    2710000
     AE017355_27      2700001    2810000
     AE017355_28      2800001    2910000
     AE017355_29      2900001    3010000
     AE017355_30      3000001    3110000
     AE017355_31      3100001    3210000
     AE017355_32      3200001    3310000
     AE017355_33      3300001    3410000
     AE017355_34      3400001    3510000
     AE017355_35      3500001    3610000
     AE017355_36      3600001    3710000
     AE017355_37      3700001    3810000
     AE017355_38      3800001    3910000
     AE017355_39      3900001    4010000
     AE017355_40      4000001    4110000
     AE017355_41      4100001    4210000
     AE017355_42      4200001    4310000
     AE017355_43      4300001    4410000
     AE017355_44      4400001    4510000
     AE017355_45      4500001    4610000
     AE017355_46      4600001    4710000
     AE017355_47      4700001    4810000
     AE017355_48      4800001    4910000
     AE017355_49      4900001    5010000
     AE017355_50      5000001    5110000
     AE017355_51      5100001    5210000
     AE017355_52      5200001    5237682
Continuation (7 of 53) of AE017355 from base 600001 (AE017355 Bacillus thuringiensis ser

  Query Match             12.4%;  Score 124.6;  DB 1;  Length 110000;
  Best Local Similarity   48.2%;  Pred. No. 3.5e-10;
  Matches  460;  Conservative    0;  Mismatches  474;  Indels   21;  Gaps    3;

Qy         32 TAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAATAGTT 91
              ||  |||| ||   |||   |||    |  ||  || |||     |   |   ||||
Db      20343 TATTTGTATTTTTATTAGCCTTTTCACTGCTTTTAAGTGCTTGTGGAAAATCAAATACGA 20284

Qy         92 CAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTA 151
               |    | ||||   ||||  |   ||||||     ||    |  | || | |   ||||
Db      20283 AAGAAGAATCAAAAGAAGATACGAAAAAAGAAATGATTCCTGTGGAACATGCAATGGGTA 20224

Qy        152 CTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATG 211
                ||  |||| |||    | || |||||||| ||  || |     ||  |  |  ||| |
Db      20223 AAACAGAAGTTCCTGCTAATCCAAAACGTGTAGTCATTTTAACAAATGAAGGTACTGAAG 20164

Qy        212 CGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTA 271
              | ||| |||  |||| ||| ||||| ||||||   |   |   |   |      |||
Db      20163 CTTTACTTGAATTAGGTGTAAAACCAGTTGGGGCTGTAAAATCTTGGACTGGTGATCCAT 20104

Qy        272 TTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAAC 331
                  |  ||   | | |||||||||   | |    |    ||| ||          |||
Db      20103 GGTATCCACATATTAAAGATAAAATGAAAGATGTAAAAGTTGTTGGTGATGAAGGACAAG 20044

Qy        332 CTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAATAGAC 391
                ||| |||||  ||||  |    | |||||||| |||||||| | | | || |   | |
Db      20043 TGAACGTAGAAACAATCGCTTCTTTAAAACCAGACTTAATTATCGGTAACAAAATGCGTC 19984

Qy        392 ACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCG 451
              || ||    | ||  |  |  ||||   |||||| |||||  |        |||  ||
Db      19983 ACGAAAAAGTGTACGAGCAACTAAAAGCAATTGCACCTACTGTGTTCTCTGAAACATTAC 19924

Qy        452 ATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAG 511
                || || |  || || ||          || ||      | ||||||| |||  |||||
Db      19923 GCGGTGAATGGAAAGATAA---------CTTCAAGTTTTATGCAAAAGCATTAAATAAAG 19873

Qy        512 AAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAG 571
              || |||| ||  ||||    ||||  ||  |||||||    ||  |||| |  |||
Db      19872 AAAAAGATGGCCAAAAGGTTTTAGCTGATTACGATAAACGTATGAAAGATTTAAAAGCTA 19813

Qy        572 AAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGC 631
              || ||  |    |   ||||||| |  | |     |                 | | | |
Db      19812 AACTAGGTGATAAGGTAAATCAAGAAATCTCTATGGTTCGCTTCATGCCTGGTGATGTTC 19753

Qy        632 TTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAG 691
               |  | |||   |  |  | |    ||||    |  |||   |||| || ||||||
Db      19752 GTATTTATCATGGTGATACATTCTCTGGTGTTATTTTAAAAGAACTTGGCTTTAAACGTC 19693

Qy        692 CATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAA 751
              |      ||  |||   | ||||| |          | | ||| | | || |    |  |
Db      19692 C------TGGCGATCAAAATAAAGATGACTTCGCAGAGCGTAACGTATCTAAAGAGCGTA 19639

Qy        752 TGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAG 811
              |     |      |     |  | |              |    | |   |   | |||
Db      19638 TTTCTGCAATGGATGGTGATGTATTATTCTACTTCACATTCGATAAAGGAAACGAGAAAA 19579

Qy        812 CAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAA 871
               | | ||| |       | |  |||||||     ||| ||||||| |||  || || ||||
Db      19578 AAGGCTCTGA------ATTAGAAAAAGAATATATAAATGATCCTCTATTTAAAAATTTAA 19525

Qy        872 ACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTG 931
              | || || ||||||      |   | |   | |||   |   | |||   | | |  |||
Db      19524 ATGCAGTAAAAAATGGTAAAGCATACAAAGTTGACGACGTTATTTGGAATACAGCTGGTG 19465

Qy        932 GTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGA 986
              || ||||  || |  |   | |   | | ||  | |   ||   | |   ||| |
Db      19464 GTGTAATGGCTGCTAACCTACTATTAGATGATATCGAAAAACGCTTTGTTAAATA 19410


RESULT 40
CP000001_06/c
WPCOMMENT
  Sequence split into 53 fragments  LOCUS CP000001 Accession CP000001
    Fragment Name        Begin       End
    CP000001_00              1    110000
    CP000001_01         100001    210000
    CP000001_02         200001    310000
    CP000001_03         300001    410000
    CP000001_04         400001    510000
    CP000001_05         500001    610000
    CP000001_06         600001    710000
    CP000001_07         700001    810000
    CP000001_08         800001    910000
    CP000001_09         900001   1010000
    CP000001_10        1000001   1110000
    CP000001_11        1100001   1210000
    CP000001_12        1200001   1310000
    CP000001_13        1300001   1410000
    CP000001_14        1400001   1510000
    CP000001_15        1500001   1610000
    CP000001_16        1600001   1710000
    CP000001_17        1700001   1810000
    CP000001_18        1800001   1910000
    CP000001_19        1900001   2010000
    CP000001_20        2000001   2110000
    CP000001_21        2100001   2210000
    CP000001_22        2200001   2310000
    CP000001_23        2300001   2410000
    CP000001_24        2400001   2510000
    CP000001_25        2500001   2610000
    CP000001_26        2600001   2710000
    CP000001_27        2700001   2810000
    CP000001_28        2800001   2910000
    CP000001_29        2900001   3010000
    CP000001_30        3000001   3110000
    CP000001_31        3100001   3210000
    CP000001_32        3200001   3310000
    CP000001_33        3300001   3410000
    CP000001_34        3400001   3510000
    CP000001_35        3500001   3610000
    CP000001_36        3600001   3710000
    CP000001_37        3700001   3810000
    CP000001_38        3800001   3910000
    CP000001_39        3900001   4010000
    CP000001_40        4000001   4110000
    CP000001_41        4100001   4210000
    CP000001_42        4200001   4310000
    CP000001_43        4300001   4410000
    CP000001_44        4400001   4510000
    CP000001_45        4500001   4610000
    CP000001_46        4600001   4710000
    CP000001_47        4700001   4810000
    CP000001_48        4800001   4910000
    CP000001_49        4900001   5010000
    CP000001_50        5000001   5110000
    CP000001_51        5100001   5210000
    CP000001_52        5200001   5300915
Continuation (7 of 53) of CP000001 from base 600001 (CP000001 Bacillus cereus ZK, comple

  Query Match             12.4%;  Score 124.6;  DB 1;  Length 110000;
  Best Local Similarity   48.2%;  Pred. No. 3.5e-10;
  Matches  460;  Conservative    0;  Mismatches  474;  Indels   21;  Gaps    3;

Qy         32 TAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAATAGTT 91
              ||  |||| ||   |||   |||    |  ||  || |||     |   |   ||||
Db      17766 TATTTGTATTTTTATTAGCCTTTTCACTACTTTTAAGTGCTTGTGGAAAATCAAATACGA 17707

Qy         92 CAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTA 151
               |    | ||||   ||||  |   ||||||     ||    |  | || | |   ||||
Db      17706 AAGAAGAATCAAAAGAAGATACGAAAAAAGAAATGATTCCTGTGGACCATGCAATGGGTA 17647

Qy        152 CTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATG 211
                ||  |||| |||    | || |||||||| ||  || |     ||  |  |  ||| |
Db      17646 AAACAGAAGTTCCTGCTAATCCAAAACGTGTAGTCATTTTAACAAATGAAGGTACTGAAG 17587

Qy        212 CGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTA 271
              | ||| |||  |||| ||| ||||| ||||||   |   |   |   |      |||
Db      17586 CTTTACTTGAATTAGGTGTGAAACCAGTTGGGGCTGTAAAATCTTGGACTGGTGATCCAT 17527
```

```
Qy        272 TTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAAC 331
                  |  ||   | | |||||||||   | |    |    ||| ||          |||
Db      17526 GGTATCCACATATTAAAGATAAAATGAAAGATGTAAAAGTTGTTGGTGATGAAGGACAAG 17467

Qy        332 CTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAATAGAC 391
                ||| |||||  ||||  |    | |||||||| |||||||| | | | || |   | |
Db      17466 TAAACGTAGAAACAATCGCTTCTTTAAAACCAGACTTAATTATCGGTAACAAAATGCGTC 17407

Qy        392 ACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCG 451
              || ||    | ||  |  |  ||||   |||||| |||||  |        |||  ||
Db      17406 ACGAAAAAGTGTACGAGCAACTAAAAGCAATTGCACCTACTGTGTTCTCTGAAACATTAC 17347

Qy        452 ATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAG 511
                || || |  || || ||          || ||      | | ||||| |||  |||||
Db      17346 GCGGTGAATGGAAAGATAA---------CTTCAAGTTTTATGCGAAAGCATTAAATAAAG 17296

Qy        512 AAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAG 571
              || |||| ||  ||||    ||||  |   |||||||    ||  |||| |  |||
Db      17295 AAAAAGATGGCCAAAAGGTTTTAGCTGCTTACGATAAACGTATGAAAGATTTAAAAGCTA 17236

Qy        572 AAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGC 631
              || ||  |    |   ||||||| |  | |     |                 | | | |
Db      17235 AACTAGGTGATAAGGTAAATCAAGAAATCTCTATGGTTCGCTTCATGCCTGGTGATGTTC 17176

Qy        632 TTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAG 691
               |  | |||   |  |  | |    ||||    |  |||   |||| || ||||||
Db      17175 GTATTTATCATGGTGATACATTCTCTGGTGTTATTTTAAAAGAACTTGGCTTTAAACGTC 17116

Qy        692 CATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAA 751
              |      ||  |||   | ||||| |     |    | | ||| | | || |    |  |
Db      17115 C------TGGCGATCAAAATAAAGATGACTTTGCAGAGCGTAACGTATCTAAAGAGCGTA 17062

Qy        752 TGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAG 811
              |     |      |     |  | |              |    | |   |   | |||
Db      17061 TTTCTGCAATGGATGGTGATGTATTATTCTACTTCACATTCGATAAAGGAAACGAGAAAA 17002

Qy        812 CAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAA 871
               | | ||| |       | |  |||||||     ||| ||||||| |||  || || ||||
Db      17001 AAGGCTCTGA------ATTAGAAAAAGAATATATAAATGATCCTCTATTTAAAAATTTAA 16948

Qy        872 ACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTG 931
              | || || ||||||      |   | |   | |||   |   | |||   | | |  |||
Db      16947 ATGCAGTAAAAAATGGTAAAGCATACAAAGTTGACGACGTTATTTGGAATACAGCTGGTG 16888

Qy        932 GTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGA 986
              || ||||  || |  |   | |   | | ||  | |   ||   | |   ||| |
Db      16887 GTGTAATGGCTGCTAACCTACTATTAGATGATATCGAAAAACGCTTTGTTAAATA 16833

Search completed: August 26, 2005, 05:54:16
Job time : 4645 secs
```

/③

```
                    GenCore version 5.1.6
             Copyright (c) 1993 - 2005  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         August 25, 2005, 23:38:01 ; Search time 644 Seconds
                                            (without alignments)
                                            9265.682 Million cell updates/sec

Title:          US-10-724-972A-2580
Perfect score:  1008
Sequence:       1 ggagtggaatcagtgagagg..........gtaaaaaagataataagtaa 1008

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters:      8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       N_Geneseq_16Dec04:*
                1:  geneseqn1980s:*
                2:  geneseqn1990s:*
                3:  geneseqn2000s:*
                4:  geneseqn2001as:*
                5:  geneseqn2001bs:*
                6:  geneseqn2002as:*
                7:  geneseqn2002bs:*
                8:  geneseqn2003as:*
                9:  geneseqn2003bs:*
                10:  geneseqn2003cs:*
                11:  geneseqn2003ds:*
                12:  geneseqn2004as:*
                13:  geneseqn2004bs:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                   %
Result             Query
  No.   Score      Match Length DB  ID                     Description
 --------------------------------------------------------------------------
     1    1008    100.0   1008 13  ADS03285               Ads03285 Staphyloc
     2     993     98.5    993  8  ACA47084               Aca47084 Prokaryot
c    3     569     56.4   2981  8  ACC48532               Acc48532 Staphyloc
     4     567     56.2    984  8  ACA19774               Aca19774 Prokaryot
     5   564.4     56.0   1014  4  AAS54519               Aas54519 Staphyloc
     6   557.4     55.3   1044  8  ACF74986               Acf74986 Staphyloc
     7   531.8     52.8    927  4  AAS51822               Aas51822 Staphyloc
c    8   336.2     33.4   2247  2  AAT80398               Aat80398 Staphyloc
c    9   336.2     33.4   2247  2  AAT83786               Aat83786 DNA encod
    10   309.8     30.7    525  8  ACA47237               Aca47237 Prokaryot
    11   269.2     26.7   2957  8  ACC48534               Acc48534 Staphyloc
    12   265.6     26.3    960  4  AAS54387               Aas54387 Staphyloc
    13   264.6     26.2    957  4  AAS51612               Aas51612 Staphyloc
    14   263.8     26.2    975  8  ACF74415               Acf74415 Staphyloc
    15   263.2     26.1    957  8  ABT15015               Abt15015 Pathogen
    16   263.2     26.1    960  8  ACA20079               Aca20079 Prokaryot
c   17   190.8     18.9   5718 12  ADF30765               Adf30765 Bacillus
    18   175.4     17.4    668  2  AAV74666               Aav74666 Staphyloc
    19   162.2     16.1    242  2  AAV76867               Aav76867 Staphyloc
    20   158.2     15.7    183  6  ABN92560               Abn92560 Staphyloc
```

```
    21   156.4     15.5    897  8  ACA42836               Aca42836 Prokaryot
    22   149.4     14.8    801  6  ABK74775               Abk74775 Bacillus
    23   143.2     14.2    972  8  ACA21545               Aca21545 Prokaryot
c   24     128     12.7    321  4  AAS49530               Aas49530 Staphyloc
c   25     128     12.7    321  8  ACA16735               Aca16735 Prokaryot
    26   125.4     12.4   2115  2  AAV74915               Aav74915 Staphyloc
    27     124     12.3    990  8  ABT15032               Abt15032 Pathogen
    28     124     12.3    999  2  AAZ22850               Aaz22850 Staphyloc
    29     124     12.3    999  2  AAZ19889               Aaz19889 Staphyloc
    30     124     12.3    999 10   ADF43555               Adf43555 Staphyloc
c   31     124     12.3   2940  8  ACC48531               Acc48531 Staphyloc
    32     124     12.3   3775  2  AAV74549               Aav74549 Staphyloc
    33   113.2     11.2    796  8  ACA22139               Aca22139 Prokaryot
    34   110.2     10.9    861  8  ACF74610               Acf74610 Staphyloc
    35   109.4     10.9   3046  4  AAH54418               Aah54418 S. epider
    36   103.2     10.2    900  8  ACA32000               Aca32000 Prokaryot
    37    96.8      9.6    909 10   ACF70890               Acf70890 Photorhab
c   38    96.8      9.6  25860 11   ADR20887               Adr20887 Photorhab
c   39    96.8      9.6 110000 10   ACF67367_39            Continuation (40 o
    40    96.8      9.6 110000 10   ACF65388_08            Continuation (9 of
c   41    96.2      9.5   1130  2  AAT83923               Aat83923 DNA encod
    42    95.2      9.4    909  4  AAS52735               Aas52735 E. coli D
    43    95.2      9.4    909  8  ACA32751               Aca32751 Prokaryot
c   44    95.2      9.4   2209  5  AAS71379               Aas71379 DNA encod
    45    95.2      9.4   2799  5  AAS77335               Aas77335 DNA encod

                              ALIGNMENTS


RESULT 1
ADS03285
ID   ADS03285 standard; DNA; 1008 BP.
XX
AC   ADS03285;
XX
DT   04-NOV-2004  (first entry)
XX
DE   Staphylococcus epidermis polynucleotide seqid 2580.
XX
KW   antibacterial; vaccine; antisense therapy; Staphylococcus epidermidis;
KW   recombinant expression vector; infection; computer readable medium;
KW   computer based system; gene; ds.
XX
OS   Staphylococcus epidermidis.
XX
PN   US2004147734-A1.
XX
PD   29-JUL-2004.
XX
PF   01-DEC-2003; 2003US-00724972.
XX
PR   08-NOV-1997;   97US-0064964P.
PR   13-AUG-1998;   98US-00134001.
PR   29-NOV-1999;   99US-00450969.
XX
PA   (DOUC/) DOUCETTE-STAMM L.
PA   (BUSH/) BUSH D.
XX
PI   Doucette-Stamm L,  Bush D;
XX
DR   WPI; 2004-580138/56.
DR   P-PSDB; ADS07057.
XX
PT   New isolated polypeptide and encoding nucleic acid derived from
PT   Staphylococcus epidermidis, useful for diagnosing, preventing and/or
PT   treating an S. epidermidis bacterial infection.
XX
PS   Claim 5; SEQ ID NO 2580; 741pp; English.
XX
CC   The invention describes an isolated nucleic acid comprising a nucleotide
CC   sequence with any of 3772 fully defined nucleotide sequences (SEQ ID NO:
CC   1-3772) and encoding an Staphylococcus epidermidis polypeptide with any
```

59-67

```
CC   of 3772 fully defined amino acid sequences (SEQ ID NO: 3772-7544) as
CC   given in the specification. Also described are: a recombinant expression
CC   vector; a cell comprising a recombinant expression vector of (1);
CC   producing an S. epidermidis polypeptide; an isolated nucleic acid
CC   comprising a nucleotide sequence of at least 8 nucleotides in length; a
CC   vaccine composition for prevention or treatment of an S. epidermidis
CC   infection, comprising a nucleic acid cited above and a carrier; treating
CC   a subject for S. epidermidis infection; a recombinant or substantially
CC   pure preparation of an S. epidermidis polypeptide or its fragment; a
CC   vaccine composition for prevention or treatment of an S. epidermidis
CC   infection; detecting the presence of a Staphylococcus nucleic acid in a
CC   sample; a computer readable medium having recorded in it the nucleotide
CC   sequences with SEQ ID NO: 1-3772 or its fragments; a computer based
CC   system for identifying fragments of the Staphylococcus genome of
CC   commercial importance; a computer based system for identifying fragments
CC   of the Staphylococcus plasmids of commercial importance; identifying
CC   commercially important nucleic acid fragments of the Staphylococcus
CC   genome and/or plasmids; and identifying an expression modulating fragment
CC   of the Staphylococcus genome and/or plasmids. The methods and
CC   compositions of the present invention are useful for the diagnosis,
CC   prevention and/or treatment of an Staphylococcal epidermidis bacterial
CC   infection. This sequence encodes a S. epidermis protein of the invention.
XX
SQ   Sequence 1008 BP; 404 A; 129 C; 188 G; 287 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 1008;  DB 13;  Length 1008;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-175;
  Matches 1008;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGAGTGGAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAGTGGAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTA 60

Qy         61 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAA 120

Qy        121 GATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGT 180

Qy        181 GTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTT 240

Qy        241 GGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGA 300

Qy        301 AAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAA 360

Qy        361 CCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAA 420

Qy        421 ATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCT 480

Qy        481 TTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAA 540

Qy        541 CACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTA 600

Qy        601 TTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGT 660

Qy        661 CAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTA 720

Qy        721 AGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAAT 780

Qy        781 CCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAA 840

Qy        841 CTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATT 900

Qy        901 TTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAA 960

Qy        961 GAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008

RESULT 2
ACA47084
ID   ACA47084 standard; DNA; 993 BP.
XX
AC   ACA47084;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Prokaryotic essential gene #28741.
XX
KW   Antisense; ds; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
XX
OS   Staphylococcus epidermidis.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,   Zamudio C,   Malone C,   Haselbeck R,   Ohlsen KL,   Zyskind JW;
PI   Wall D,   Trawick JD,   Carr GJ,   Yamamoto R,   Forsyth RA,   Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   P-PSDB; ABU43214.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 14; SEQ ID NO 34954; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential genes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 993 BP; 398 A; 128 C; 183 G; 284 T; 0 U; 0 Other;

  Query Match             98.5%;  Score 993;  DB 8;  Length 993;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-172;
  Matches  993;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           13 GTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCA 72
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1 GTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCA 60

Qy           73 GCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAA 132
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           61 GCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAA 120

Qy          133 ATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTT 192
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          121 ATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTT 180

Qy          193 GAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGAT 252
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 GAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGAT 240

Qy          253 GATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCT 312
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 GATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCT 300

Qy          313 GTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATT 372
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 GTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATT 360

Qy          373 ATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACG 432
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          361 ATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACG 420

Qy          433 ATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATT 492
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 ATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATT 480

Qy          493 TCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAA 552
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 TCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAA 540

Qy          553 ATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTA 612
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 ATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTA 600

Qy          613 GCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGT 672
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          601 GCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGT 660

Qy          673 CAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTT 732
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          661 CAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTT 720

Qy          733 AAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATG 792
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          721 AAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATG 780

Qy          793 TTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGAT 852
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          781 TTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGAT 840

Qy          853 CCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGAC 912
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          841 CCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGAC 900

Qy          913 TTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAA 972
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          901 TTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAA 960

Qy          973 TTATCTAAGAAAGATAGTAAAAAAGATAATAAG 1005
                |||||||||||||||||||||||||||||||||
Db          961 TTATCTAAGAAAGATAGTAAAAAAGATAATAAG 993


RESULT 3
ACC48532/c
ID   ACC48532 standard; DNA; 2981 BP.
XX
AC   ACC48532;
XX
DT   11-AUG-2003  (first entry)
XX
DE   Staphylococcal surface-exposed immunogenic polypeptide DNA.
XX
KW   Surface-exposed immunogenic polypeptide; SEIP; siderophore; receptor;
KW   antibacterial; vaccine; gene; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   CDS             complement(968. .1951)
FT                   /*tag=  a
FT                   /product= "SEIP"
XX
PN   WO2003020875-A2.
XX
PD   13-MAR-2003.
XX
PF   17-JUN-2002; 2002WO-US019224.
XX
PR   17-JUN-2001; 2001US-0298975P.
XX
PA   (DSQU-) D-SQUARED BIOTECHNOLOGIES INC.
XX
PI   Scott DL;
XX
DR   WPI; 2003-300870/29.
DR   P-PSDB; ABR41865.
XX
PT   Novel surface-exposed immunogenic polypeptide of Staphylococcus aureus
PT   containing receptors for siderophores or iron-binding ligands, useful for
PT   producing antibodies effective against Staphylococci infection.
XX
PS   Claim 4; Page 53; 62pp; English.
XX
CC   The present sequence is that of DNA encoding novel Staphylococcus aureus
CC   surface-exposed immunogenic polypeptide (SEIP) D2 SA03. To isolate SEIP
CC   genes, an expression library of S. aureus genomic DNA was screened with
CC   anti-SEIP antisera. DNA was isolated from clones expressing SEIPs, and
CC   amino acid sequences were deduced and analysed for conserved and/or
CC   functional domains. The SEIP encoded by the present DNA sequence has
CC   sequence homology with the siderophore family of periplasmic binding
CC   proteins. The invention provides methods for the identification,
CC   production and recovery of SEIPs. The SEIPs can be used individually, or
CC   in combination, to produce anti-staphylococcal antibodies useful in
CC   passive or active immunisation strategies to prevent or contain
CC   staphylococcal infection. They can also be used to develop diagnostic
CC   assays
XX
SQ   Sequence 2981 BP; 988 A; 559 C; 386 G; 1048 T; 0 U; 0 Other;

  Query Match             56.4%;  Score 569;  DB 8;  Length 2981;
  Best Local Similarity   73.3%;  Pred. No. 6.8e-95;
  Matches  728;  Conservative    0;  Mismatches  265;  Indels    0;  Gaps    0;

Qy          7 GAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCA 66
              | | |  |||||||||| |||||| ||| ||| || | || ||| |  |||  ||| ||  |
Db       1987 GGAACGATGAGAGGTCTAAAAACTTTTAGTATATTGGGATTAATAGTTGCCTTACTTTTA 1928

Qy         67 ACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGA 126
                |||||| ||||| ||||    |      |    ||   ||   ||| | ||||||   
Db       1927 GTTGCAGCTTGTGGTAATACGGATAATTCAAGTAAAAAAGAATCATCAACTAAAGATACT 1868

Qy        127 GTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTT 186
               ||    | ||  | ||| | |||||    |||||||||||| |  | ||||||| | |||
Db       1867 ATTTCGGTAAAAGATGAAAATGGTACAGTAAAAGTACCTAAAGATGCAAAACGTATCGTT 1808

Qy        187 GTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATA 246
              ||  | ||||| ||||||| ||||| ||||  || ||||| |||||||||| ||||| || 
Db       1807 GTATTAGAGTACTCATTTGCAGATGCATTAGCAGCATTAGACGTTAAACCAGTTGGTATT 1748

Qy        247 GCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATAC 306
              || ||||||   || ||||| ||||| ||||||||| | ||||| |||||||||  | || 
Db       1747 GCTGATGATGGTAAGAAAAAACGTATCATTAAACCAGTTAGAGAAAAAATTGGGGATTAT 1688

Qy        307 ACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGAT 366
              ||||||||||| |||||||| || || |||||||| |||||| |||||| | ||||| |||
Db       1687 ACTTCTGTAGGTACACGTAAACAGCCAAACTTAGAGGAAATTAGTAAATTAAAACCGGAT 1628

Qy        367 TTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCT 426
              |||||||| |||||||  | |||||| |||||||||| |||||||| ||||| |||||||| 
Db       1627 TTAATTATCGCTGATAGCAGTAGACATAAAGGTATTAATAAAGAATTAAACAAAATTGCA 1568

Qy        427 CCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAA 486
              || ||  |   | | || |||| ||||||||| || ||  |||||||| || | || |||
Db       1567 CCAACATTATCATTAAAGAGTTTTGATGGAGACTACAAACAAAATATTAATTCGTTCAAA 1508

Qy        487 ACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGAT 546
              |||||| | |||||||||  ||||||| |||||||  |||||||||  | |  || || |||
Db       1507 ACAATTGCTAAAGCTTTAAATAAAGAAAAAGAAGGCGAAAAACGTCTTGCTGAGCATGAT 1448

Qy        547 AAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCT 606
              ||   | |  |  | ||||||| | ||||| |   | |||| ||||||||| ||  | || 
Db       1447 AAATTAATCAATAAGTATAAAGATGAAATTAAATTTGATAGAAATCAAAAAGTGCTTCCA 1388

Qy        607 GCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTC 666
              ||||||| |||||| | ||||  | || ||||||| | | || |||||||| ||||| 
Db       1387 GCAGTAGTTGCTAAAGCTGGTTTATTAGCACATCCAAACTATTCATATGTTGGACAATTT 1328

Qy        667 CTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAG 726
               |||   |||||||| |||||| | |||||||||||| |||| || |||||||||||||| 
Db       1327 TTAAACGAACTAGGATTTAAAAATGCATTAAGTGACGATGTAACAAAAGGTTTAAGTAAA 1268

Qy        727 TATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAG 786
              ||| | ||||||||||||||||||| |  ||||||||  |||| || |  | ||||| |||
Db       1267 TATTTGAAAGGACCTTACTTACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAG 1208

Qy        787 CGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAA 846
              |||||| |||| |||||| |  | || |      |     ||  | | ||  |  || ||
Db       1207 CGTATGATCATTATGACAGATCATGCTAAAAAAGATTCTGCTGAATTCAAGAAGTTACAA 1148

Qy        847 AAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGAC 906
               ||||| |   |||||| || || || || |||||| | || || |||||| | |||
Db       1147 GAAGATGCAACATGGAAAAAGTTGAATGCAGTTAAAAATAATCGCGTGGATATTGTTGAC 1088

Qy        907 CGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTT 966
              |||||  | ||||||||||| ||||| |||||||||||| ||||||||||| |||||||||
Db       1087 CGTGATGTTTGGGCAAGATCTCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTT 1028

Qy        967 GTTGAATTATCTAAGAAAGATAGTAAAAAAGAT 999
              ||||||||||| || |||||    ||  ||| |
Db       1027 GTTGAATTATCAAAAAAAGAACAAAAGTAAGGT 995


RESULT 4
ACA19774
ID   ACA19774 standard; DNA; 984 BP.
XX
AC   ACA19774;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Prokaryotic essential gene #1431.
XX
KW   Antisense; ds; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
XX
OS   Staphylococcus aureus.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,  Zamudio C,  Malone C,  Haselbeck R,  Ohlsen KL,  Zyskind JW;
PI   Wall D,  Trawick JD,  Carr GJ,  Yamamoto R,  Forsyth RA,  Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   P-PSDB; ABU15904.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 14; SEQ ID NO 7644; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential genes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 984 BP; 399 A; 123 C; 181 G; 281 T; 0 U; 0 Other;

  Query Match             56.2%;  Score 567;  DB 8;  Length 984;
  Best Local Similarity   73.6%;  Pred. No. 1.5e-94;
  Matches  723;  Conservative    0;  Mismatches  260;  Indels    0;  Gaps    0;

Qy          14 TGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAG 73
               |||||||| |||||| ||| ||| || | || ||| |  |||  ||| ||  |  |||||
Db           2 TGAGAGGTCTAAAAACTTTTAGTATATTGGGATTAATAGTTGCCTTACTTTTAGTTGCAG 61

Qy          74 CATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAA 133
               | ||||| ||||    |      |    ||   ||   ||| | ||||||    ||
Db          62 CTTGTGGTAATACGGATAATTCAAGTAAAAAAGAATCATCAACTAAAGATACTATTTCGG 121

Qy         134 TCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTG 193
               | ||  | ||| | |||||    |||||||||||| |  | |||||| | |||||  | |
Db         122 TAAAAGATGAAAATGGTACAGTAAAAGTACCTAAAGATGCAAAACGTATCGTTGTATTAG 181

Qy         194 AGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATG 253
               |||| |||||||  ||||| ||||  || ||||| |||||||| ||||| || || ||||
Db         182 AGTACTCATTTGCAGATGCATTAGCAGCATTAGACGTTAAACCAGTTGGTATTGCTGATG 241

Qy         254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
               ||   || ||||| ||||| ||||||||| | ||||| ||||||||  | || |||||||
Db         242 ATGGTAAGAAAAAACGTATCATTAAACCAGTTAGAGAAAAAATTGGGGATTATACTTCTG 301

Qy         314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
               |||| |||||||| || || |||||||| ||||| |||||| | ||||| ||||||||||
Db         302 TAGGTACACGTAAACAGCCAAACTTAGAGGAAATTAGTAAATTAAAACCGGATTTAATTA 361

Qy         374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
               | |||||||  | |||||| ||||||||| ||||||| ||||| |||||||| || ||
Db         362 TCGCTGATAGCAGTAGACATAAAGGTATTAATAAAGAATTAAACAAAATTGCACCAACAT 421

Qy         434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
               |   | | || ||||| |||||||| || ||  |||||||| || | || |||||||||
Db         422 TATCATTAAAGAGTTTTGATGGAGACTACAAACAAAATATTAATTCGTTCAAAACAATTG 481

Qy         494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
               | |||||||||  ||||||| |||||||  |||||||  | |  || || |||||   ||
Db         482 CTAAAGCTTTAAATAAAGAAAAAGAAGGCGAAAAACGTCTTGCTGAGCATGATAAATTAA 541

Qy         554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
               |  |  | |||||| | ||||| |   | |||| ||||||||| ||  | || |||||||
Db         542 TCAATAAGTATAAAGATGAAATTAAATTTGATAGAAATCAAAAAGTGCTTCCAGCAGTAG 601

Qy         614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
                ||||||| | |||||  | || ||||||| | | || |||||||| |||||  |||
Db         602 TTGCTAAAGCTGGTTTATTAGCACATCCAAACTATTCATATGTTGGACAATTTTTAAACG 661

Qy         674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
               ||||||| |||||| | ||||||||||| ||||| || |||||||||||||| ||| | |
Db         662 AACTAGGATTTAAAAATGCATTAAGTGACGATGTAACAAAAGGTTTAAGTAAATATTTGA 721

Qy         734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
               ||||||||||||||||| |  ||||||||  |||| || |  | ||||| |||||||||
Db         722 AAGGACCTTACTTACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGA 781

Qy         794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
               |||| |||||| |  | || |      |     ||  | | ||  |  || || |||||
Db         782 TCATTATGACAGATCATGCTAAAAAAGATTCTGCTGAATTCAAGAAGTTACAAGAAGATG 841
```

```
Qy         854 CTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACT 913
               |   |||||| || || || || || |||||| | || || |||||| | ||||||||
Db         842 CAACATGGAAAAAGTTGAATGCAGTTAAAAATAATCGCGTGGATATTGTTGACCGTGATG 901

Qy         914 TATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAAT 973
               | ||||||||||| ||||| ||||||||||| ||||||||||| ||||||||||||||||
Db         902 TTTGGGCAAGATCTCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTTGTTGAAT 961

Qy         974 TATCTAAGAAAGATAGTAAAAAA 996
               |||| || |||||    ||  ||
Db         962 TATCAAAAAAAGAACAAAAGTAA 984

RESULT 5
AAS54519
ID   AAS54519 standard; DNA; 1014 BP.
XX
AC   AAS54519;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus DNA for cellular proliferation protein #831.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
KW   antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US009180.
XX
PR   21-MAR-2000; 2000US-0191078P.
PR   23-MAY-2000; 2000US-0206848P.
PR   26-MAY-2000; 2000US-0207727P.
PR   23-OCT-2000; 2000US-0242578P.
PR   27-NOV-2000; 2000US-0253625P.
PR   22-DEC-2000; 2000US-0257931P.
PR   16-FEB-2001; 2001US-0269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU36660.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids.
XX
PS   Claim 27; SEQ ID NO 8156; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the genes,
CC   their use in the discovery of novel antibiotics, the essential genes
CC   themselves and the encoded proteins. The prokaryotes used are Escherichia
CC   coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC   Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC   useful for the identification of potential new targets for antibiotic
CC   development. The antisense nucleic acids can also be used to identify
CC   proteins used in proliferation, to express these proteins, and to obtain
CC   antibodies capable of binding to the expressed proteins. The proteins can
CC   be used to screen compounds in rational drug discovery programmes. The
CC   antisense nucleic acid sequence is also useful to screen for homologous
CC   nucleic acids which are required for cell proliferation in a wide variety
CC   of organisms. The present sequence encodes an essential prokaryotic
CC   cellular proliferation protein. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
```

```
XX
SQ   Sequence 1014 BP; 410 A; 126 C; 187 G; 291 T; 0 U; 0 Other;

  Query Match             56.0%;  Score 564.4;  DB 4;  Length 1014;
  Best Local Similarity   73.1%;  Pred. No. 4.4e-94;
  Matches  724;  Conservative    0;  Mismatches  266;  Indels    0;  Gaps    0;

Qy           7 GAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCA 66
               | | |  |||||||| |||||| ||| ||| || | || ||| |  |||  ||| ||  |
Db          25 GGAACGATGAGAGGTCTAAAAACTTTTAGTATATTGGGATTAATAGTTGCCTTATTTTTA 84

Qy          67 ACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGA 126
                 |||||| ||||| ||||    |      |    ||   ||   ||| | ||||||
Db          85 GTTGCAGCTTGTGGTAATACGGATAATTCAAGTAAAAAAGAATCATCAACTAAAGATACT 144

Qy         127 GTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTT 186
                ||    | ||  | ||| | |||||    |||||||||||| |  | |||||| | |||
Db         145 ATTTCGGTAAAAGATGAAAATGGTACAGTAAAAGTACCTAAAGATGCAAAACGTATCGTT 204

Qy         187 GTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATA 246
               ||  | ||||| |||||||  ||||| ||||  || ||||| |||||||| ||||| ||
Db         205 GTATTAGAGTACTCATTTGCAGATGCATTAGCAGCATTAGACGTTAAACCAGTTGGTATT 264

Qy         247 GCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATAC 306
               || ||||||   || ||||| ||||| ||||||||| | ||||| |||||||| || ||
Db         265 GCTGATGATGGTAAGAAAAAAACGTATCATTAAACCAGTTAGAGAAAAAATTGGGAATTAT 324

Qy         307 ACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGAT 366
               ||||||||||| |||||||| || || |||||||| ||||| |||||| | ||||| |||
Db         325 ACTTCTGTAGGTACACGTAAACAGCCAAACTTAGAGGAAATTAGTAAATTAAAACCGGAT 384

Qy         367 TTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCT 426
               |||||||| |||||||  | |||||| ||||||||| ||||||| ||||| ||||||||
Db         385 TTAATTATCGCTGATAGCAGTAGACATAAAGGTATTAATAAAGAATTAAACAAAATTGCA 444

Qy         427 CCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAA 486
               || ||  |   | | || ||||| |||||||| || ||  |||| ||| || | || |||
Db         445 CCAACATTATCATTAAAGAGTTTTGATGGAGACTACAAACAAAACATTAATTCGTTCAAA 504

Qy         487 ACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGAT 546
               |||||| | |||||||||  ||||||| |||||||  |||| ||  | |  ||||| |||
Db         505 ACAATTGCTAAAGCTTTAAATAAAGAAAAAGAAGGCGAAAAGCGTCTTGCTGAACATGAT 564

Qy         547 AAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCT 606
               ||   |||  || | |||||| | ||||| |   | |||| ||||||||| ||  | ||
Db         565 AAATTAATCAAAAAGTATAAAGATGAAATTAAGTTTGATAGAAATCAAAAAGTGCTTCCA 624

Qy         607 GCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTC 666
               ||||| | ||||||| | |||||  | || ||||||| | | || |||||||| |||||
Db         625 GCAGTTGTTGCTAAAGCTGGTTTATTAGCACATCCAAACTATTCATATGTTGGACAATTT 684

Qy         667 CTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAG 726
                |||   |||| || |||||| | ||||||||||||||||| || ||||||||||||||
Db         685 TTAAACGAACTTGGATTTAAAAATGCATTAAGTGATGATGTAACAAAAGGTTTAAGTAAA 744

Qy         727 TATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAG 786
               ||  | |||||||||||||||||| |  | ||||||  |||| || |  | ||||||||
Db         745 TACTTGAAAGGACCTTACTTACAATTAGATACTGAACATTTAGCTGACTTAAATCCTGAA 804

Qy         787 CGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAA 846
               || ||| | || |||||| | || || |      |     ||  | | ||  |  || ||
Db         805 CGCATGATTATTATGACAGATAATGCTAAAAAAGATTCTGCTGAATTCAAGAAGTTACAA 864

Qy         847 AAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGAC 906
                |||||||    ||||| || || ||||| || |||||| | || || |||||| | |||
Db         865 GAAGATCCAACTTGGAAAAAGTTGAACGCAGTTAAAAATAATCGCGTGGATATTGTTGAC 924

Qy         907 CGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTT 966
               |||||  | ||||||||||| ||||| ||||||||||| ||||||||||| |||||||||
Db         925 CGTGATGTTTGGGCAAGATCTCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTT 984

Qy         967 GTTGAATTATCTAAGAAACATAGTAAAAAA 996
               ||||||||||| || ||| |    ||  ||
```

```
Db         985 GTTGAATTATCAAAAAAAGAACAAAAGTAA 1014

RESULT 6
ACF74986
ID   ACF74986 standard; DNA; 1044 BP.
XX
AC   ACF74986;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Staphylococcus aureus DNA #2666.
XX
KW   Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW   enzymatic assay; antibiotic target; gene; ds.
XX
OS   Staphylococcus aureus.
XX
PN   WO200294868-A2.
XX
PD   28-NOV-2002.
XX
PF   27-MAR-2002; 2002WO-IB002637.
XX
PR   27-MAR-2001; 2001GB-00007661.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Masignani V,  Mora M,  Scarselli M;
XX
DR   WPI; 2003-120786/11.
DR   P-PSDB; ABM73426.
XX
PT   New Staphylococcus aureus protein, useful as a vaccine for treating or
PT   preventing Staphylococcal infection, specifically an infection caused by
PT   S. aureus, e.g. sepsis.
XX
PS   Claim 6; SEQ ID NO 5331; 49pp; English.
XX
CC   The invention relates to novel genes and encoded proteins from
CC   Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC   nucleic acid encoding the protein, or an antibody to the protein, is
CC   useful as a pharmaceutical, particularly as a vaccine for treating or
CC   preventing infection due to Staphylococcus bacteria, specifically an
CC   infection caused by S. aureus. The composition is particularly useful for
CC   treating or preventing sepsis in a patient. The composition can also be
CC   used for diagnostics. The protein is also used in an assay for enzymatic
CC   studies and as a target for antibiotics. This sequence represents one of
CC   the novel S. aureus genes of the invention
XX
SQ   Sequence 1044 BP; 425 A; 128 C; 191 G; 300 T; 0 U; 0 Other;

  Query Match             55.3%;  Score 557.4;  DB 8;  Length 1044;
  Best Local Similarity   73.5%;  Pred. No. 8.5e-93;
  Matches  711;  Conservative    0;  Mismatches  256;  Indels    0;  Gaps    0;

Qy          20 GTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTG 79
               | |  |||  ||| ||| || | || ||| |  |||  ||| ||  |  |||||| ||||
Db          71 GGTCTAAACCTTTTAGTATATTGGGATTAATAGTTGCCTTACTTTTAGTTGCAGCTTGTG 130

Qy          80 GAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGC 139
               | ||||    |      |    ||   ||   ||| | ||||||    ||    | ||
Db         131 GTAATACGGATAATTCAAGTAAAAAAGAATCATCAACTAAAGATACTATTTCGGTAAAAG 190

Qy         140 ACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATT 199
               | ||| | |||||    |||||||||||| |  | |||||| | |||||  | ||||| |
Db         191 ATGAAAATGGTACAGTAAAAGTACCTAAAGATGCAAAACGTATCGTTGTATTAGAGTACT 250

Qy         200 CATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACA 259
               ||||||  ||||| ||||  || ||||| |||||||| ||||| || || ||||||   |
Db         251 CATTTGCAGATGCATTAGCAGCATTAGACGTTAAACCAGTTGGTATTGCTGATGATGGTA 310

Qy         260 AAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAA 319
```

```
              | ||||| ||||| ||||||||| | ||||| ||||||||  | || ||||||||||| |
Db        311 AGAAAAAACGTATCATTAAACCAGTTAGAGAAAAAAATTGGGATTATACTTCTGTAGGTA 370

Qy        320 CACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTG 379
              ||||||| || || |||||||||||||| |||||| | ||||| |||||||||||| ||||
Db        371 CACGTAAACAGCCAAACTTAGAAGAAATTAGTAAATTAAAACCGGATTTAATTATCGCTG 430

Qy        380 ATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAAC 439
              |||  | |||||| ||||||||| ||||||| ||||| |||||||| || ||  |    |
Db        431 ATAGCAGTAGACATAAAGGTATTAATAAAGAATTAAACAAAATTGCACCAACATTATCAT 490

Qy        440 TGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAG 499
              | || ||||| |||||||| || ||  |||||||| || | || ||||||||| | ||||
Db        491 TAAAGAGTTTTGATGGAGACTACAAACAAAATATTAATTCGTTCAAAACAATTGCTAAAG 550

Qy        500 CTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAG 559
              |||||  ||||||| |||||||  |||||||  | |  || || |||||   |||  |
Db        551 CTTTAAATAAAGAAAAAGAAGGCGAAAAACGTCTTGCTGAGCATGATAAATTAATCAATA 610

Qy        560 AATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTA 619
              | |||||| | ||||| |   | |||| ||||||||| ||  | || ||||||| |||||
Db        611 AGTATAAAGATGAAATTAAATTTGATAGAAATCAAAAAGTGCTTCCAGCAGTAGTTGCTA 670

Qy        620 AATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAG 679
              || | |||||  | || ||||||| | | || |||||||| |||||  |||   ||||||
Db        671 AAGCTGGTTTATTAGCACATCCAAACTATTCATATGTTGGACAATTTTTAAACGAACTAG 730

Qy        680 GTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGAC 739
              | |||||| | ||||||||||| ||||| || |||||||||||||| ||| | |||||||
Db        731 GATTTAAAAATGCATTAAGTGACGATGTAACAAAAGGTTTAAGTAAATATTTGAAAGGAC 790

Qy        740 CTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAA 799
              ||||||||||| |  ||||||||  |||| || |  | ||||| ||||||||| |||| |
Db        791 CTTACTTACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGATCATTA 850

Qy        800 TGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTAT 859
              ||||| |  | || |      |     ||  | | ||  |  || || ||||| |   ||
Db        851 TGACAGATCATGCTAAAAAAGATTCTGCTGAATTCAAGAAGTTACAAGAAGATGCAACAT 910

Qy        860 GGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGG 919
              |||| || || || || || |||||| | || || ||||||| | ||||||||  | ||||
Db        911 GGAAAAAGTTGAATGCAGTTAAAAATAATCGCGTGGATATTGTTGACCGTGATGTTTGGG 970

Qy        920 CAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTA 979
              ||||||| ||||| ||||||||||| ||||||||||| |||||||||||||||||||| |
Db        971 CAAGATCTCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTTGTTGAATTATCAA 1030

Qy        980 AGAAAGA 986
              | |||||
Db       1031 AAAAAGA 1037


RESULT 7
AAS51822
ID   AAS51822 standard; DNA; 927 BP.
XX
AC   AAS51822;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus DNA for cellular proliferation protein #239.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
KW   antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US009180.
XX
PR   21-MAR-2000; 2000US-0191078P.
PR   23-MAY-2000; 2000US-0206848P.
PR   26-MAY-2000; 2000US-0207727P.
PR   23-OCT-2000; 2000US-0242578P.
PR   27-NOV-2000; 2000US-0253625P.
PR   22-DEC-2000; 2000US-0257931P.
PR   16-FEB-2001; 2001US-0269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU33963.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids.
XX
PS   Claim 27; SEQ ID NO 4404; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the genes,
CC   their use in the discovery of novel antibiotics, the essential genes
CC   themselves and the encoded proteins. The prokaryotes used are Escherichia
CC   coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC   Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC   useful for the identification of potential new targets for antibiotic
CC   development. The antisense nucleic acids can also be used to identify
CC   proteins used in proliferation, to express these proteins, and to obtain
CC   antibodies capable of binding to the expressed proteins. The proteins can
CC   be used to screen compounds in rational drug discovery programmes. The
CC   antisense nucleic acid sequence is also useful to screen for homologous
CC   nucleic acids which are required for cell proliferation in a wide variety
CC   of organisms. The present sequence encodes an essential prokaryotic
CC   cellular proliferation protein. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 927 BP; 370 A; 120 C; 169 G; 268 T; 0 U; 0 Other;

  Query Match             52.8%;  Score 531.8;  DB 4;  Length 927;
  Best Local Similarity   73.4%;  Pred. No. 4e-88;
  Matches  680;  Conservative    0;  Mismatches  247;  Indels    0;  Gaps    0;

Qy         46 TTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGT 105
              ||| |  |||  ||| ||  |  |||||| ||||| ||||    |      |    ||
Db          1 TTAATAGTTGCCTTATTTTTAGTTGCAGCTTGTGGTAATACGGATAATTCAAGTAAAAAA 60

Qy        106 AAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCT 165
               ||   ||| | ||||||    ||    | ||  | ||| | |||||    |||||||||
Db         61 GAATCATCAACTAAAGATACTATTTCGGTAAAAGATGAAAATGGTACAGTAAAAGTACCT 120

Qy        166 AAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTA 225
              ||| |  | |||||| | |||||  | ||||| |||||||  ||||| ||||  || |||
Db        121 AAAGATGCAAAACGTATCGTTGTATTAGAGTACTCATTTGCAGATGCATTAGCAGCATTA 180

Qy        226 GATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTA 285
              || ||||||||| ||||| || || ||||||   || ||||| ||||| ||||||||| |
Db        181 GACGTTAAACCAGTTGGTATTGCTGATGATGGTAAGAAAAAACGTATCATTAAACCAGTT 240

Qy        286 AGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAA 345
              ||||| |||||||| || || ||||||||||| |||||||| || || |||||||| |||
Db        241 AGAGAAAAAATTGGGAATTATACTTCTGTAGGTACACGTAAACAGCCAAACTTAGAGGAA 300

Qy        346 ATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTAT 405
              || |||||| | ||||| ||||||||||| |||||||  | |||||| ||||||||| ||
Db        301 ATTAGTAAATTAAAACCGGATTTAATTATCGCTGATAGCAGTAGACATAAAGGTATTAAT 360

Qy        406 AAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAAT 465
              ||||| ||||| |||||||| || ||  |   | | || ||||| |||||||| || ||
Db        361 AAAGAATTAAACAAAATTGCACCAACATTATCATTAAAGAGTTTTGATGGAGACTACAAA 420
```

```
Qy        466 GAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAA 525
               |||| ||| || | || ||||||||| | |||||||||  ||||||| |||||||  ||
Db        421 CAAAACATTAATTCGTTCAAAACAATTGCTAAAGCTTTAAATAAAGAAAAAGAAGGCGAA 480

Qy        526 AAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGAT 585
              || ||  | |  ||||| |||||   |||  || | |||||| | ||||| |   | |||
Db        481 AAGCGTCTTGCTGAACATGATAAATTAATCAAAAAGTATAAAGATGAAATTAAGTTTGAT 540

Qy        586 AAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGC 645
              | ||||||||| ||  | || ||||| | ||||||| | |||||  | || ||||||| |
Db        541 AGAAATCAAAAAGTGCTTCCAGCAGTTGTTGCTAAAGCTGGTTTATTAGCACATCCAAAC 600

Qy        646 AACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGAT 705
               | || |||||||| |||||  |||   |||| || |||||| | |||||||||||||||
Db        601 TATTCATATGTTGGACAATTTTTAAACGAACTTGGATTTAAAAATGCATTAAGTGATGAT 660

Qy        706 GTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACT 765
              || || |||||||||||||| ||  | |||||||||||||||||| |  | ||||||  |
Db        661 GTAACAAAAGGTTTAAGTAAATACTTGAAAGGACCTTACTTACAATTAGATACTGAACAT 720

Qy        766 TTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAA 825
              ||| || |  | |||||||| || ||| | || |||||| | || || |      |
Db        721 TTAGCTGACTTAAATCCTGAACGCATGATTATTATGACAGATAATGCTAAAAAAGATTCT 780

Qy        826 CCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAAT 885
               ||  | | ||  |  || || |||||||    ||||| || || ||||| || ||||||
Db        781 GCTGAATTCAAGAAGTTACAAGAAGATCCAACTTGGAAAAAGTTGAACGCAGTTAAAAAT 840

Qy        886 CAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCA 945
               | || || |||||| | ||||||||  | ||||||||||| ||||| |||||||||||
Db        841 AATCGCGTGGATATTGTTGACCGTGATGTTTGGGCAAGATCTCGTGGCTTAATTTCTTCT 900

Qy        946 GAAGAAATGGCAAAAGAACTTGTTGAA 972
              ||||||||||| |||||||||||||||
Db        901 GAAGAAATGGCTAAAGAACTTGTTGAA 927


RESULT 8
AAT80398/c
ID   AAT80398 standard; cDNA; 2247 BP.
XX
AC   AAT80398;
XX
DT   14-APR-1998  (first entry)
XX
DE   Staphylococcus aureus Gene #4 encoding cDNA sequence 1.
XX
KW   Staphylococcus aureus WCUH 29; antagonist; antibacterial; immunogen;
KW   vaccine; disease; protection; isolation; ss.
XX
OS   Staphylococcus aureus.
XX
PN   WO9731114-A2.
XX
PD   28-AUG-1997.
XX
PF   25-FEB-1997;   97WO-GB000524.
XX
PR   26-FEB-1996;   96GB-00004045.
XX
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Burnham MKR,  Hodgson JE;
XX
DR   WPI; 1997-435166/40.
XX
PT   New Staphylococcus aureus polynucleotide and polypeptide(s) - for
PT   isolating antagonist of the polypeptide(s) useful as anti-bacterials.
XX
PS   Claim 4; Page 30-31; 117pp; English.
XX
CC   The present sequence encodes a novel polypeptide, which is optionally
CC   expressed in NCIMB 40771. The polypeptide, and polynucleotide encoding
CC   it, are derived from Staphylococcus aureus. Cells expressing ligands
CC   binding the polypeptide can be used to isolated candidate compounds that
CC   bind and inhibit the activity of the polypeptide. Such compounds can be
CC   used as anti-bacterial compounds. The polypeptide may also be used as an
CC   immunogen to vaccinate an animal for protection against Staphylococcus
CC   aureus caused disease
XX
SQ   Sequence 2247 BP; 704 A; 397 C; 307 G; 802 T; 0 U; 37 Other;

  Query Match             33.4%;  Score 336.2;  DB 2;  Length 2247;
  Best Local Similarity   70.6%;  Pred. No. 2.5e-52;
  Matches  580;  Conservative    0;  Mismatches  229;  Indels   13;  Gaps   10;

Qy          1 GGAGTGGAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTA 60
              ||||||| |    |||||||| |||||| ||| ||| || | || ||| |  |||  |||
Db        831 GGAGTGGGACNGATGAGAGGTCTAAAAACTTTTAGTATATTGGGATTAATAGTTGCCTTA 772

Qy         61 ATTGCAACTGCAGCATGTGGAAATA---ATAGTTCAAGTAACTCAAGTAAAGAGTCATCA 117
               ||  |  |||||| ||||| ||||   ||| |||||||||   ||| ||  |   |  |
Db        771 CTTTTAGTTGCAGCTTGTGGTAATACGGATAATTCAAGTAAAAGAAG-AATCATCAACTA 713

Qy        118 AAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAA 177
              ||      | ||    | ||  | ||| | |||||    |||||||||||| |  |  ||
Db        712 AAGATACTATTTCGGGTAAAAGATGAAAATGGTACAGTAAAAGTACCTAAAGATGCACAA 653

Qy        178 CGTGTTGTTGTTCTTGAGTATTCATTTG--TTGATGCGTTAGTTGC-TTTAGATGTTAAA 234
              ||| | |||||  | ||||| |||||||    | ||| |||   || | |||  |||| |
Db        652 CGTATCGTTGTATTAGAGTACTCATTTGCAAGGTTGCATTANCAGCATATAGCCGTTANA 593

Qy        235 CC-TGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAA 293
              ||  ||||| || || ||||||   || ||||| ||||| | ||| ||| ||||||| ||
Db        592 CCAAGTTGGTATTGCTGATGATGGTAAGAAAAAACGTATCA-TAACCCAGTAAGAGA-AA 535

Qy        294 AATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAA 353
              ||||||     | ||||||||||| |||||||| ||| | ||||||  ||| || |||||
Db        534 AATTGGGGGTAATACTTCTGTAGGTACACGTAA-CAAGCNAACTTAAGAGANATTAGTAA 476

Qy        354 ACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTT 413
              | | || || |  ||||||||  ||||||  | |||||| || |||||| ||||||| ||
Db        475 ATTAAANCCGGGATTAATTATCNCTGATAGCAGTAGACATAAGGGTATTAATAAAGAATT 416

Qy        414 AAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATAT 473
              |    | ||||| || ||  |   | | || ||||| |||||||| || |   | |||||
Db        415 AGCCCACATTGCACC-ACATTATCATTAAAGAGTTTTGATGGAGACTACACCCACAATAT 357

Qy        474 TGATGCTTTTAA-AACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCT 532
              | || | || || ||||||| | |||||||||  ||||||  |||||||    |||||
Db        356 TAATTCGTTCAACAACAATTGCTAAAGCTTTAAATAAAGANAAAGAAGGCGNNAAACGTC 297

Qy        533 TAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATC 592
              | |  || || |||||   |||  |  | |||||| | ||||| |   | |||| |||||
Db        296 TTGCTGAGCATGATAAATTAATCAATAAGTATAAAGATGAAATTAAATTTGATAGAAATC 237

Qy        593 AAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTT 652
              | || ||  | || ||||||| ||||||| | |||||  | || ||||||  | | || |
Db        236 ACAAAGTGCTTCCAGCAGTAGTTGCTAAAGCTGGTTTATTAGCACATCCACACTATTCAT 177

Qy        653 ATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTA 712
              ||||||| |||||  |||   | ||||| |||||| | ||||||||||| ||||| ||
Db        176 ATGTTGGACAATTTTTAAACGANCTAGGATTTAAAAATGCATTAAGTGACGATGTAACAC 117

Qy        713 AAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTC 772
              ||||||||||||| ||| | |||||||||||||||||| |  ||||||||  |||| ||
Db        116 AAGGTTTAAGTAAATATTTGAAAGGACCTTACTTACAATTAGACACTGAACATTTAGCTG 57

Qy        773 AAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAA 814
              |  | ||||| ||||||||| |||| |||||| |  | || |
Db         56 ATTTAAATCCAGAGCGTATGATCATTATGACAGATCATGCTA 15


RESULT 9
AAT83786/c
```

```
ID   AAT83786 standard; DNA; 2247 BP.
XX
AC   AAT83786;
XX
DT   16-JUL-1998  (first entry)
XX
DE   DNA encoding 3 Staphylococcus aureus proteins of unknown function.
XX
KW   Staphylococcus aureus protein; ribozyme; antisense sequence; control;
KW   Staphylococcal gene; regulatory element; bacterial gene expression;
KW   vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;
KW   toxic shock syndrome; ss.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   CDS             complement(1212..1445)
FT                   /*tag=  a
FT   CDS             complement(1456..1587)
FT                   /*tag=  b
FT   CDS             complement(1705..2016)
FT                   /*tag=  c
XX
PN   WO9730070-A1.
XX
PD   21-AUG-1997.
XX
PF   19-FEB-1997;   97WO-US002318.
XX
PR   20-FEB-1996;   96US-0011888P.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Black MT,  Burnham MK,  Hodgson JE,  Knowles DJC,  Nicholas RO;
PI   Pratt JM,  Reichard RW,  Rosenberg M,  Ward JM;
XX
DR   WPI; 1997-424969/39.
DR   P-PSDB; AAW27819, AAW27820, AAW27821.
XX
PT   Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used to
PT   isolate antimicrobial compounds, and in vaccines against S. aureus
PT   infection.
XX
PS   Claim 9; Page 672-673; 989pp; English.
XX
CC   The present sequence encodes 3 Staphylococcus aureus proteins of unknown
CC   function. The present sequence was isolated from a library of clones of
CC   S. aureus WCUH 29 in Escherichia coli. The DNA sequence can be used in
CC   the construction of ribozymes and antisense sequences to control the
CC   expression of Staphylococcal genes. The DNA sequence is also useful as a
CC   source of regulatory elements for the control of bacterial gene
CC   expression. The encoded protein may be used to produce vaccines to enable
CC   a host to produce specific antibodies with antibacterial action. These
CC   vaccines and antibodies would protect a host against invasion by S.
CC   aureus, and conditions relating to Staphylococcal infection, e.g.
CC   Staphylococcal food poisoning, scaled skin syndrome, and toxic shock
CC   syndrome
XX
SQ   Sequence 2247 BP; 704 A; 397 C; 307 G; 802 T; 0 U; 37 Other;

  Query Match             33.4%;  Score 336.2;  DB 2;  Length 2247;
  Best Local Similarity   70.6%;  Pred. No. 2.5e-52;
  Matches  580;  Conservative    0;  Mismatches  229;  Indels   13;  Gaps   10;

Qy           1 GGAGTGGAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTA 60
               ||||||| |    |||||||| |||||| ||| || ||| || | || ||| |  |||  |||
Db         831 GGAGTGGGACNGATGAGAGGTCTAAAAACTTTTAGTATATTGGGATTAATAGTTGCCTTA 772

Qy          61 ATTGCAACTGCAGCATGTGGAAATA---ATAGTTCAAGTAACTCAAGTAAAGAGTCATCA 117
                ||  |  ||||| ||||| ||||   ||| |||||||||   ||| ||  |   |  |
Db         771 CTTTTAGTTGCAGCTTGTGGTAATACGGATAATTCAAGTAAAAGAAG-AATCATCAACTA 713

Qy         118 AAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAA 177
               ||      | ||    | ||  | ||| | |||||    ||||||||||| | |  |  ||
Db         712 AAGATACTATTTCGGGTAAAAGATGAAAATGGTACAGTAAAAGTACCTAAAGATGCACAA 653

Qy         178 CGTGTTGTTGTTCTTGAGTATTCATTTG--TTGATGCGTTAGTTGC-TTTAGATGTTAAA 234
               ||| | |||||  | ||||| |||||||    | ||| |||   || | |||  |||| |
Db         652 CGTATCGTTGTATTAGAGTACTCATTTGCAAGGTTGCATTANCAGCATATAGCCGTTANA 593

Qy         235 CC-TGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAA 293
               ||  ||||| || || ||||||   || ||||| ||||| | ||| ||| ||||||| ||
Db         592 CCAAGTTGGTATTGCTGATGATGGTAAGAAAAAACGTATCA-TAACCCAGTAAGAGA-AA 535

Qy         294 AATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAA 353
               ||||||      | ||||||||||| |||||||| ||| | ||||||  ||| || |||||
Db         534 AATTGGGGGTAATACTTCTGTAGGTACACGTAA-CAAGCNAACTTAAGAGANATTAGTAA 476

Qy         354 ACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTT 413
               | | || || |  |||||||||  ||||||   | |||||| || |||||| ||||||| ||
Db         475 ATTAAANCCGGGATTAATTATCNCTGATAGCAGTAGACATAAGGGTATTAATAAAGAATT 416

Qy         414 AAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATAT 473
               |    | ||||| || ||  |   | | || ||||| |||||||| || |   | |||||
Db         415 AGCCCACATTGCACC-ACATTATCATTAAAGAGTTTTGATGGAGACTACACCCACAATAT 357

Qy         474 TGATGCTTTTAA-AACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCT 532
               | || | || || ||||||| | |||||||||  ||||||  |||||||    |||||
Db         356 TAATTCGTTCAACAACAATTGCTAAAGCTTTAAATAAAGANAAAGAAGGCGNNAAACGTC 297

Qy         533 TAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATC 592
               | |  || || |||||   |||  |  | ||||| | ||||| |   | |||| |||||
Db         296 TTGCTGAGCATGATAAATTAATCAATAAGTATAAAGATGAAATTAAATTTGATAGAAATC 237

Qy         593 AAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTT 652
               | || ||  | || ||||||| ||||||| | |||||  | || ||||||  | | || |
Db         236 ACAAAGTGCTTCCAGCAGTAGTTGCTAAAGCTGGTTTATTAGCACATCCACACTATTCAT 177

Qy         653 ATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTA 712
               ||||||| ||||||  |||   | ||||| |||||| | ||||||||||| ||||| ||
Db         176 ATGTTGGACAATTTTTAAACGANCTAGGATTTAAAAATGCATTAAGTGACGATGTAACAC 117

Qy         713 AAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTC 772
               ||||||||||||| ||| | ||||||||||||||||||| |  ||||||||  |||| ||
Db         116 AAGGTTTAAGTAAATATTTGAAAGGACCTTACTTACAATTAGACACTGAACATTTAGCTG 57

Qy         773 AAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAA 814
               |  | |||||  ||||||||| |||| ||||| |  | ||  |
Db          56 ATTTAAATCCAGAGCGTATGATCATTATGACAGATCATGCTA 15


RESULT 10
ACA47237
ID   ACA47237 standard; DNA; 525 BP.
XX
AC   ACA47237;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Prokaryotic essential gene #28894.
XX
KW   Antisense; ds; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
XX
OS   Staphylococcus haemolyticus.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,  Zamudio C,  Malone C,  Haselbeck R,  Ohlsen KL,  Zyskind JW;
PI   Wall D,  Trawick JD,  Carr GJ,  Yamamoto R,  Forsyth RA,  Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   P-PSDB; ABU43367.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 14; SEQ ID NO 35107; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential genes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 525 BP; 200 A; 66 C; 96 G; 163 T; 0 U; 0 Other;

  Query Match             30.7%;  Score 309.8;  DB 8;  Length 525;
  Best Local Similarity   76.3%;  Pred. No. 1.6e-47;
  Matches  399;  Conservative    0;  Mismatches  112;  Indels   12;  Gaps    1;

Qy         14 TGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAG 73
              ||| ||| || ||  ||  | || || | |  ||||| ||||  ||| ||  |  | | |
Db          2 TGAAAGGCTTTAAGTTTGCAGGTATAGTCGCATTATTATTTGCATTAGTTTTAGTTACTG 61

Qy         74 CATGTGGAAATAATAGTTCAAGTAACTCAAG------------TAAAGAGTCATCAAAAG 121
              | ||||| |||  |||    | |   ||| |            ||||  |||||| ||||
Db         62 CTTGTGGTAATGTTAGCAATAATGGATCAGGTGATTCAGGTAATAAATCGTCATCGAAAG 121

Qy        122 ATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTG 181
              ||    || |||| || ||||||  ||| || || ||||| |||||| |  |||||||||
Db        122 ATTCTATTAAAATTAAACACGAATTAGGCACAACTAAAGTTCCTAAAGATGCTAAACGTG 181

Qy        182 TTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTG 241
              | || |   | || | ||||||||| ||||| ||||  ||  || |||||||||||||||
Db        182 TCGTAGCGTTGGAATTTTCATTTGTAGATGCTTTAGCGGCACTAAATGTTAAACCTGTTG 241

Qy        242 GGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAA 301
              || | || ||||| ||||||  ||||||||||||||||||||||| ||| |||||||||
Db        242 GGGTTGCTGATGACAACAAACCAAATCGTATTATTAAACCATTAAAAGAAAAAATTGGAG 301

Qy        302 AATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAAC 361
              | || |  ||||| ||  | ||||| ||||| ||||||||||||||||||||| | ||||
Db        302 ATTATAAATCTGTTGGTGCTCGTAAACAACCCAACTTAGAAGAAATCAGTAAATTAAAAC 361

Qy        362 CAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAA 421
              |||||||||||||||| |||| ||| ||||||||||||||||| ||||| |||| |||||
Db        362 CAGATTTAATTATTGCAGATAGTAACAGACACAAAGGTATTTACAAAGAATTAAGTAAAA 421

Qy        422 TTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTT 481
              ||||||| || |||||| | |||||||| ||||||||||||||||| ||||| |||||||
Db        422 TTGCTCCAACAATTGAATTAAAAAGTTTTGATGGAGATTATAATGATAATATAGATGCTT 481

Qy        482 TTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAA 524
              ||||||||||| | |||||||||  |||||| || | |||| |
Db        482 TTAAAACAATTGCCAAAGCTTTAAATAAAGATGATGTAGGTCA 524


RESULT 11
ACC48534
ID   ACC48534 standard; DNA; 2957 BP.
XX
AC   ACC48534;
XX
DT   11-AUG-2003  (first entry)
XX
DE   Staphylococcal surface-exposed immunogenic polypeptide DNA.
XX
KW   Surface-exposed immunogenic polypeptide; SEIP; siderophore; receptor;
KW   antibacterial; vaccine; gene; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   CDS             983. .1960
FT                   /*tag=  a
FT                   /partial
FT                   /product= "SEIP"
FT                   /note= "No start codon"
XX
PN   WO2003020875-A2.
XX
PD   13-MAR-2003.
XX
PF   17-JUN-2002; 2002WO-US019224.
XX
PR   17-JUN-2001; 2001US-0298975P.
XX
PA   (DSQU-) D-SQUARED BIOTECHNOLOGIES INC.
XX
PI   Scott DL;
XX
DR   WPI; 2003-300870/29.
DR   P-PSDB; ABR41867.
XX
PT   Novel surface-exposed immunogenic polypeptide of Staphylococcus aureus
PT   containing receptors for siderophores or iron-binding ligands, useful for
PT   producing antibodies effective against Staphylococci infection.
XX
PS   Claim 4; Page 55-56; 62pp; English.
XX
CC   The present sequence is that of DNA encoding novel Staphylococcus aureus
CC   surface-exposed immunogenic polypeptide (SEIP) D2 SA05. To isolate SEIP
CC   genes, an expression library of S. aureus genomic DNA was screened with
CC   anti-SEIP antisera. DNA was isolated from clones expressing SEIPs, and
CC   amino acid sequences were deduced and analysed for conserved and/or
CC   functional domains. The SEIP encoded by the present DNA sequence has
CC   sequence homology with the siderophore family of periplasmic binding
CC   proteins. The invention provides methods for the identification,
CC   production and recovery of SEIPs. The SEIPs can be used individually, or
CC   in combination, to produce anti-staphylococcal antibodies useful in
CC   passive or active immunisation strategies to prevent or contain
CC   staphylococcal infection. They can also be used to develop diagnostic
CC   assays
XX
SQ   Sequence 2957 BP; 1160 A; 381 C; 512 G; 904 T; 0 U; 0 Other;

  Query Match             26.7%;  Score 269.2;  DB 8;  Length 2957;
  Best Local Similarity   55.4%;  Pred. No. 4.6e-40;
  Matches  545;  Conservative    0;  Mismatches  433;  Indels    6;  Gaps    1;

Qy         17 GAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCAT 76
              |||||  | |  |    ||    || |    |||  |||| || ||   ||    ||
Db        990 GAGGTAAATATGTGAATAGGAATATCGTTAAATTAGTTGTGTTTATGCTAATCTTAGTTG 1049

Qy         77 GTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCA 136
                | |       |||   || |   |  ||  ||      ||  ||         ||| |
Db       1050 TAGCAGTAGCGGGTTGTGGTCAAAAAGATACTGAAGAGAAAACTGAAATGACGACAATAA 1109

Qy        137 AGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGT 196
              |  | |||  ||| |||   ||| |    ||| | |||||||||||||||||  | || |
Db       1110 AAGATGAATTAGGAACTGAAAAAATTAAGAAAAATCCTAAACGTGTTGTTGTATTAGAAT 1169

Qy        197 ATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
              ||   |||| ||||   ||||  || |||||| | ||||||||||| || || ||||||
Db       1170 ATAGTTTTGCTGATTATTTAGCAGCATTAGATATGAAACCTGTTGGTATTGCAGATGATG 1229

Qy        257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
               ||  |  ||   ||| |  ||  || |||||||||| |||||   |||    || || |
Db       1230 GCAGCACTAAAAATATAACAAAGTCAGTAAGAGATAAGATTGGGGCATATGAATCGGTTG 1289

Qy        317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
              || |  |   |||||| ||  | ||||  || |||||| | ||||| ||||| || ||||
Db       1290 GATCTAGACCGCAACCGAATATGGAAGTGATAAGTAAATTAAAACCGGATTTGATCATTG 1349

Qy        377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
              | |||  ||  ||||| ||    ||  |   ||| || |  ||||||||||| || ||
Db       1350 CAGATGTTAGCAGACATAAGAAAATCAAATCAGAATTGAGCAAAATTGCTCCGACAATCA 1409

Qy        437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
                 |    ||       ||||||||||||| ||||||||| || ||||||||| |  | |
Db       1410 TGTTAGTTAGCGGTACGGGAGATTATAATGCAAATATTGAAGCATTTAAAACAGTCGCTA 1469

Qy        497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
              ||||  |||| |||||  |||||||  | || ||  | ||| | || ||||| | | | |
Db       1470 AAGCAGTAGGCAAAGAGAAAGAAGGCGAGAAGCGTCTGGAAAAGCATGATAAAATATTAG 1529

Qy        557 AAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTG 616
                ||   || |||  ||||      |  ||      ||   | |||  |    |
Db       1530 CGGAGATTAGAAAGAAAATTGAACAGAGTACGTTAAAATCTGCATTTGCATTCGGTATCT 1589

Qy        617 CTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAAC 676
              | | | ||||| || ||  | ||       |  | | | | || |||||| ||| | ||
Db       1590 CAAGAGCAGGTATGTTTATTAATAATGAAGATACATTTATGGGACAATTCTTAATTAAAA 1649

Qy        677 TAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAG 736
              | ||| || ||   | | | |   | ||    ||||    | |  || |          |
Db       1650 TGGGTATTCAACCTGAAGTCACAAAAGACAAAACTACGCATGTTGGTGAACGCAAGGGTG 1709

Qy        737 GACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCA 796
              | |||||  || |  | || | ||||    |  |  |  | |||||  |   |||| |
Db       1710 GTCCTTATATATATTTAAATAATGAAGAACTTGCCAATATCAATCCAAAAGTTATGATTT 1769

Qy        797 TAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTG 856
              ||   ||  ||  |  ||     ||  |         ||   || |     |   | | |
Db       1770 TAGCCACTGACGGAAAAACGGACAAAAA------TAGAACGAAATTCATTGATCCTGCAG 1823

Qy        857 TATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTAT 916
              | |||||   |||||| ||||||||| || |    ||| ||    | |||||  |    |
Db       1824 TTTGGAAATCATTAAAAGCTGTGAAAGATAACAAAGTTTATGACGTTGACCGAAATAAGT 1883

Qy        917 GGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTAT 976
              ||   | |||| | ||  | ||  |    |||   |||||| ||||  | |   || |
Db       1884 GGTTGAAATCAAGGGGGATTATCGCAAGTGAAAGTATGGCAGAAGATTTAGAAAAAATTG 1943

Qy        977 CTAAGAAAGATAGTAAAAAAGATA 1000
              |  | ||||  |   ||||| | |
Db       1944 CAGAAAAAGCAAAATAAAAATACA 1967

RESULT 12
AAS54387
ID   AAS54387 standard; DNA; 960 BP.
XX
AC   AAS54387;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus DNA for cellular proliferation protein #699.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
KW   antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US009180.
XX
PR   21-MAR-2000; 2000US-0191078P.
PR   23-MAY-2000; 2000US-0206848P.
PR   26-MAY-2000; 2000US-0207727P.
PR   23-OCT-2000; 2000US-0242578P.
PR   27-NOV-2000; 2000US-0253625P.
PR   22-DEC-2000; 2000US-0257931P.
PR   16-FEB-2001; 2001US-0269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU36528.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids.
XX
PS   Claim 27; SEQ ID NO 8024; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the genes,
CC   their use in the discovery of novel antibiotics, the essential genes
CC   themselves and the encoded proteins. The prokaryotes used are Escherichia
CC   coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC   Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC   useful for the identification of potential new targets for antibiotic
CC   development. The antisense nucleic acids can also be used to identify
CC   proteins used in proliferation, to express these proteins, and to obtain
CC   antibodies capable of binding to the expressed proteins. The proteins can
CC   be used to screen compounds in rational drug discovery programmes. The
CC   antisense nucleic acid sequence is also useful to screen for homologous
CC   nucleic acids which are required for cell proliferation in a wide variety
CC   of organisms. The present sequence encodes an essential prokaryotic
CC   cellular proliferation protein. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 960 BP; 385 A; 115 C; 209 G; 251 T; 0 U; 0 Other;

  Query Match             26.3%;  Score 265.6;  DB 4;  Length 960;
  Best Local Similarity   57.8%;  Pred. No. 1.9e-39;
  Matches  494;  Conservative    0;  Mismatches  354;  Indels    6;  Gaps    1;

Qy        132 AATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCT 191
              ||| ||  | |||  ||| |||   ||| |    ||| | |||||||||||||||||  |
Db        105 AATAAAAGATGAATTAGGAACTGAAAAAATTAAGAAAAATCCTAAACGTGTTGTTGTATT 164

Qy        192 TGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGA 251
               || |||   |||| ||||   ||||  || |||||| | ||||||||||| || || ||
Db        165 AGAATATAGTTTTGCTGATTATTTAGCAGCATTAGATATGAAACCTGTTGGTATTGCAGA 224

Qy        252 TGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTC 311
              ||||  ||  |  ||   ||| |  ||  || ||||||||||  ||||   |||    ||
Db        225 TGATGGCAGCAGTAAAAATATAACAAAGTCAGTAAGAGATAAGGTTGGGGCATATGAATC 284

Qy        312 TGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAAT 371
               || ||| |  |   ||||||| ||  | ||||  || |||||| | ||||| ||||| ||
Db        285 GGTTGGATCTAGACCGCAACCGAATATGGAAGTGATAAGTAAATTAAAACCGGATTTGAT 344

Qy        372 TATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTAC 431
               || || |||  ||  ||||| ||    ||  |   ||| ||||  ||||||||||| ||
Db        345 CATAGCAGATGTTAGCAGACATAAGAAAATCAAATCAGAATTAAGCAAAATTGCTCCGAC 404

Qy        432 GATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAAT 491
               ||     |    ||       ||||||||||||| ||||||||||||| ||||||||| |
Db        405 AATCATGTTAGTTAGCGGTACGGGAGATTATAATGCAAATATTGATGCATTTAAAACAGT 464

Qy        492 TTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAA 551
                | |||||  |||| |||||  |||||||  | |||||  | ||| | || ||||| |
Db        465 CGCTAAAGCAGTAGGCAAAGAGAAAGAAGGCGAGAAACGTCTGGAAAAGCATGATAAAAT 524

Qy        552 AATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGT 611
              | | |  ||   || |||  ||||      |  ||      |||  | |||  |    |
Db        525 ATTAGCGGAGATTAGAAAGAAAATTGAACAGAGTACGTTAAAAACTGCATTTGCATTCGG 584

Qy        612 AGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAG 671
                   | | | ||||| || ||  | ||       |  | | | | || |||||| |||
Db        585 TATCTCAAGAGCAGGTATGTTTATTAATAATGAAGATACATTTATGGGACAATTCTTAAT 644

Qy        672 TCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCT 731
              | || | ||| || ||   | | | |   | ||    ||||    | |  || |
Db        645 TAAAATGGGTATTCAACCTGAAGTCACAAAAGACAAAACTACGCATGTTGGTGAACGCAA 704

Qy        732 TAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTAT 791
                  || |||||  |  |  | || | ||||    |  |  |  | |||||  |   |||
Db        705 GGGTGGTCCTTATATTTATTTAAATAATGAAGAACTTGCCAATATCAATCCAAAAGTTAT 764

Qy        792 GTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGA 851
              | |  ||   ||  ||  |  ||     ||  |  `      ||   || |     |
Db        765 GATTTTAGCCACTGACGGAAAAACGGACAAAAATAG------AACGAAATTCATTGATCC 818

Qy        852 TCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGA 911
              | | || |||||   |||||| |||||||||| || |    ||| ||    | |||||  |
Db        819 TGCAGTTTGGAAATCATTAAAAGCTGTGAAAGATAACAAAGTTTATGACGTTGACCGAAA 878

Qy        912 CTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGA 971
                  |||   | |||| | ||| | ||  |    |||   |||||| ||||  | |   |
Db        879 TAAGTGGTTGAAATCAAGGGGTATTATCGCAAGTGAAAGTATGGCAGAAGATTTAGAAAA 938

Qy        972 ATTATCTAAGAAAG 985
              | |  |  | ||||
Db        939 AATTGCAGAAAAAG 952

RESULT 13
AAS51612
ID   AAS51612 standard; DNA; 957 BP.
XX
AC   AAS51612;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus DNA for cellular proliferation protein #29.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
KW   antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US009180.
XX
PR   21-MAR-2000; 2000US-0191078P.
PR   23-MAY-2000; 2000US-0206848P.
PR   26-MAY-2000; 2000US-0207727P.
PR   23-OCT-2000; 2000US-0242578P.
PR   27-NOV-2000; 2000US-0253625P.
PR   22-DEC-2000; 2000US-0257931P.
PR   16-FEB-2001; 2001US-0269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU33753.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids.
XX
PS   Claim 27; SEQ ID NO 4194; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the genes,
CC   their use in the discovery of novel antibiotics, the essential genes
CC   themselves and the encoded proteins. The prokaryotes used are Escherichia
CC   coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC   Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC   useful for the identification of potential new targets for antibiotic
CC   development. The antisense nucleic acids can also be used to identify
CC   proteins used in proliferation, to express these proteins, and to obtain
CC   antibodies capable of binding to the expressed proteins. The proteins can
CC   be used to screen compounds in rational drug discovery programmes. The
CC   antisense nucleic acid sequence is also useful to screen for homologous
CC   nucleic acids which are required for cell proliferation in a wide variety
CC   of organisms. The present sequence encodes an essential prokaryotic
CC   cellular proliferation protein. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 957 BP; 382 A; 114 C; 209 G; 252 T; 0 U; 0 Other;

  Query Match             26.2%;  Score 264.6;  DB 4;  Length 957;
  Best Local Similarity   57.8%;  Pred. No. 3e-39;
  Matches  493;  Conservative    0;  Mismatches  354;  Indels    6;  Gaps    1;

Qy        132 AATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCT 191
              ||| ||  | |||  ||| |||   ||| |    ||| | |||||||||||||||||  |
Db        111 AATAAAAGATGAATTAGGAACTGAAAAAATTAAGAAAAATCCTAAACGTGTTGTTGTATT 170

Qy        192 TGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGA 251
               || |||   |||| ||||   ||||  || |||||| | ||||||||||| || || ||
Db        171 AGAATATAGTTTTGCTGATTATTTAGCAGCATTAGATATGAAACCTGTTGGTATTGCAGA 230

Qy        252 TGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTC 311
              ||||  ||  |  ||   ||| |  ||  || ||||||||||  ||||   |||    ||
Db        231 TGATGGCAGCAGTAAAAATATAACAAAGTCAGTAAGAGATAAGGTTGGGGCATATGAATC 290

Qy        312 TGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAAT 371
               || ||| |  |   ||||||| ||  | ||||  || |||||| | ||||| ||||| ||
Db        291 GGTTGGATCTAGACCGCAACCGAATATGGAAGTGATAAGTAAATTAAAACCGGATTTGAT 350

Qy        372 TATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTAC 431
               || || |||  ||  ||||| ||    ||  |   ||| ||||  ||||||||||| ||
Db        351 CATAGCAGATGTTAGCAGACATAAGAAAATCAAATCAGAATTAAGCAAAATTGCTCCGAC 410
```

```
Qy        432 GATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAAT 491
               ||     |    ||       ||||||||||||| |||||||||||| ||||||||| |
Db        411 AATCATGTTAGTTAGCGGTACGGGAGATTATAATGCAAATATTGATGCATTTAAAACAGT 470

Qy        492 TTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAA 551
                | |||||  |||| |||||  |||||||  | |||||  | ||| | || ||||| |
Db        471 CGCTAAAGCAGTAGGCAAAGAGAAAGAAGGCGAGAAACGTCTGGAAAAGCATGATAAAAT 530

Qy        552 AATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGT 611
              | | |  ||   || |||  ||||      |  ||      |||  | |||  |    |
Db        531 ATTAGCGGAGATTAGAAAGAAAATTGAACAGAGTACGTTAAAAACTGCATTTGCATTCGG 590

Qy        612 AGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAG 671
                   | | | ||||| || ||  | ||       |  | | | | || |||||| |||
Db        591 TATCTCAAGAGCAGGTATGTTTATTAATAATGAAGATACATTTATGGGACAATTCTTAAT 650

Qy        672 TCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCT 731
              | || | ||| || ||   | | | |   | ||    ||||    | |  || |
Db        651 TAAAATGGGTATTCAACCTGAAGTCACAAAAGACAAAACTACGCATGTTGGTGAACGCAA 710

Qy        732 TAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTAT 791
                  || |||||  |  |  | || | ||||    |  |  |  | |||||  |   |||
Db        711 GGGTGGTCCTTATATTTATTTAAATAATGAAGAACTTGCCAATATCAATCCAAAAGTTAT 770

Qy        792 GTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGA 851
              | |  ||   ||  ||  |  ||     ||  |         ||   || |     |
Db        771 GATTTTAGCCACTGACGGAAAAACGGACAAAAATAG------AACGAAATTCATTGATCC 824

Qy        852 TCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGA 911
              | | || |||||   |||||| ||||||||| || |    ||| ||    | |||||  |
Db        825 TGCAGTTTGGAAATCATTAAAAGCTGTGAAAGATAACAAAGTTTATGACGTTGACCGAAA 884

Qy        912 CTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGA 971
                  |||   | |||| | ||| | ||  |    |||   |||||| ||||  | |   |
Db        885 TAAGTGGTTGAAATCAAGGGGTATTATCGCAAGTGAAAGTATGGCAGAAGATTTAGAAAA 944

Qy        972 ATTATCTAAGAAA 984
              | |  |  | |||
Db        945 AATTGCAGAAAAA 957

RESULT 14
ACF74415
ID   ACF74415 standard; DNA; 975 BP.
XX
AC   ACF74415;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Staphylococcus aureus DNA #2095.
XX
KW   Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW   enzymatic assay; antibiotic target; gene; ds.
XX
OS   Staphylococcus aureus.
XX
PN   WO200294868-A2.
XX
PD   28-NOV-2002.
XX
PF   27-MAR-2002; 2002WO-IB002637.
XX
PR   27-MAR-2001; 2001GB-00007661.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Masignani V,  Mora M,  Scarselli M;
XX
DR   WPI; 2003-120786/11.
DR   P-PSDB; ABM72855.
XX
PT   New Staphylococcus aureus protein, useful as a vaccine for treating or
PT   preventing Staphylococcal infection, specifically an infection caused by
PT   S. aureus, e.g. sepsis.
XX
PS   Claim 6; SEQ ID NO 4189; 49pp; English.
XX
CC   The invention relates to novel genes and encoded proteins from
CC   Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC   nucleic acid encoding the protein, or an antibody to the protein, is
CC   useful as a pharmaceutical, particularly as a vaccine for treating or
CC   preventing infection due to Staphylococcus bacteria, specifically an
CC   infection caused by S. aureus. The composition is particularly useful for
CC   treating or preventing sepsis in a patient. The composition can also be
CC   used for diagnostics. The protein is also used in an assay for enzymatic
CC   studies and as a target for antibiotics. This sequence represents one of
CC   the novel S. aureus genes of the invention
XX
SQ   Sequence 975 BP; 387 A; 116 C; 216 G; 256 T; 0 U; 0 Other;

  Query Match             26.2%;  Score 263.8;  DB 8;  Length 975;
  Best Local Similarity   55.3%;  Pred. No. 4.2e-39;
  Matches  536;  Conservative    0;  Mismatches  427;  Indels    6;  Gaps    1;

Qy         17 GAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCAT 76
              |||||  |    |    ||    || |    | |  |||| || ||   || |  ||
Db          8 GAGGTAAATGTGTGAATAGGAATATCGTTAAACTAGTTGTGTTCATGCTAATTTTAGTTG 67

Qy         77 GTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCA 136
                | |       |||   || |   |  ||  ||      ||  ||         ||| |
Db         68 TAGCAGTAGCGGGTTGTGGTCAAAAAGATACTGAAGAGAAAACTGAAATGACGACAATAA 127

Qy        137 AGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGT 196
              |  | |||  ||| |||   ||| |    ||| | |||||||||||||||||  | || |
Db        128 AAGATGAATTAGGAACTGAAAAAATTAAGAAAAATCCTAAACGTGTTGTTGTATTAGAAT 187

Qy        197 ATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
              ||   |||| ||||   ||||  || |||||| | ||||||||||| || || ||||||
Db        188 ATAGTTTTGCTGATTATTTAGCAGCATTAGATATGAAACCTGTTGGTATTGCAGATGATG 247

Qy        257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
               ||  |  ||   ||| |  ||  || |||||||||| |||||   |||    || || |
Db        248 GCAGCACTAAAAATATAACAAAGTCAGTAAGAGATAAGATTGGGGCATATGAATCGGTTG 307

Qy        317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
              || |  |   |||||| ||  | ||||  || |||||| | ||||| ||||| || ||||
Db        308 GATCTAGACCGCAACCGAATATGGAAGTGATAAGTAAATTAAAACCGGATTTGATCATTG 367

Qy        377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
              | |||  ||  ||||| ||    ||  |   ||| || |  ||||||||||| || ||
Db        368 CAGATGTTAGCAGACATAAGAAAATCAAATCAGAATTGAGCAAAATTGCTCCGACAATCA 427

Qy        437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
                 |    ||       ||||||||||||| ||||||||| || ||||||||| |  | |
Db        428 TGTTAGTTAGCGGTACGGGAGATTATAATGCAAATATTGAAGCATTTAAAACAGTCGCTA 487

Qy        497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
              ||||  |||| |||||  |||||||  | || ||  | ||| | || ||||| | | | |
Db        488 AAGCAGTAGGCAAAGAGAAAGAAGGCGAGAAGCGTCTGGAAAAGCATGATAAAATATTAG 547

Qy        557 AAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTG 616
                ||   || |||  ||||      |  ||      ||   | |||  |    |
Db        548 CGGAGATTAGAAAGAAAATTGAACAGAGTACGTTAAAATCTGCATTTGCATTCGGTATCT 607

Qy        617 CTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAAC 676
              | | | ||||| || ||  | ||       |  | | | | || |||||| ||| | ||
Db        608 CAAGAGCAGGTATGTTTATTAATAATGAAGATACATTTATGGGACAATTCTTAATTAAAA 667

Qy        677 TAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAG 736
              | ||| || ||   | | | |   | ||    ||||    | |  || |          |
Db        668 TGGGTATTCAACCTGAAGTCACAAAAGACAAAACTACGCATGTTGGTGAACGCAAGGGTG 727

Qy        737 GACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCA 796
              | |||||  || |  | || | ||||    |  |  |  | |||||  |   |||| |
Db        728 GTCCTTATATATATTTAAATAATGAAGAACTTGCCAATATCAATCCAAAAGTTATGATTT 787
```

```
Qy         797 TAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTG 856
               ||   ||  ||  |  ||     ||  |         ||   || |     |   | | |
Db         788 TAGCCACTGACGGAAAAACGGACAAAAA------TAGAACGAAATTCATTGATCCTGCAG 841

Qy         857 TATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTAT 916
               | |||||   |||||| ||||||||| || |    ||| ||    | |||||  |    |
Db         842 TTTGGAAATCATTAAAAGCTGTGAAAGATAACAAAGTTTATGACGTTGACCGAAATAAGT 901

Qy         917 GGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTAT 976
               ||   | |||| | ||  | ||  |    |||   |||||| ||||  | |   || |
Db         902 GGTTGAAATCAAGGGGGATTATCGCAAGTGAAAGTATGGCAGAAGATTTAGAAAAAATTG 961

Qy         977 CTAAGAAAG 985
               |  | ||||
Db         962 CAGAAAAAG 970


RESULT 15
ABT15015
ID   ABT15015 standard; DNA; 957 BP.
XX
AC   ABT15015;
XX
DT   06-MAR-2003  (first entry)
XX
DE   Pathogen specific antigen related staphylococcal DNA SEQ ID No 301.
XX
KW   Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;
KW   hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
KW   auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
KW   autoimmune disease; HIV; hepatitis; gene; ds.
XX
OS   Staphylococcus sp.
XX
PN   WO200259148-A2.
XX
PD   01-AUG-2002.
XX
PF   21-JAN-2002; 2002WO-EP000546.
XX
PR   26-JAN-2001; 2001AT-00000130.
XX
PA   (CIST-) CISTEM BIOTECHNOLOGIES GMBH.
XX
PI   Meinke A,  Nagy E,  Von Ahsen U,  Klade C,  Henics T,  Zauner W;
PI   Minh DB,  Vytvytska O,  Etz H,  Dryla A,  Weichhart T,  Hafner M;
PI   Tempelmaier B;
XX
DR   WPI; 2003-075410/07.
XX
PT   Identifying, isolating and producing hyperimmune serum-reactive antigens
PT   from a pathogen, for preparing vaccine or medicament for treating or
PT   preventing e.g. staphylococcal infections, comprises providing antibody
PT   preparation.
XX
PS   Example 7; Page 203; 252pp; English.
XX
CC   The invention relates to a novel method for identifying, isolating and
CC   producing hyperimmune serum-reactive antigens from a pathogen, tumour,
CC   allergen, a tissue or host prone to auto-immunity, where the antigens are
CC   used in a vaccine, comprises providing antibody preparation from a plasma
CC   pool of a type of animal, or individual sera with antibodies against the
CC   specific pathogen, tumour, allergen, tissue or host prone to auto-
CC   immunity. The hyperimmune serum-reactive antigens comprising any of the
CC   62 sequences of 53-2261 amino acids fully defined in the specification,
CC   or their hyperimmune fragments are useful for the manufacture of a
CC   pharmaceutical preparation, particularly a vaccine against staphylococcal
CC   infections or colonisation against S. aureus or S. epidermidis. The
CC   preparation of antibodies is useful for the manufacture of a medicament
CC   for treating or preventing staphylococcal infections or colonisation
CC   against S. aureus or S. epidermidis. The antibody preparations may also
CC   be used for diagnostic and imaging purposes. Other conditions that can be
CC   treated include cancer, autoimmune diseases or infections caused by viral
CC   (e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This
CC   polynucleotide sequence represents staphylococcal DNA relating to the
CC   method for identifying and producing pathogen specific antigens of the
CC   invention
XX
SQ   Sequence 957 BP; 383 A; 116 C; 210 G; 248 T; 0 U; 0 Other;

  Query Match             26.1%;  Score 263.2;  DB 8;  Length 957;
  Best Local Similarity   55.9%;  Pred. No. 5.4e-39;
  Matches  524;  Conservative    0;  Mismatches  408;  Indels    6;  Gaps    1;

Qy          48 ATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAA 107
               | |  |||| || ||   || |  ||     | |       |||   || |   |  ||
Db          21 ACTAGTTGTGTTCATGCTAATTTTAGTTGTAGCAGTAGCGGGTTGTGGTCAAAAAGATAC 80

Qy         108 AGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAA 167
                ||      ||  ||         ||| ||  | |||  ||| |||   ||| |    ||
Db          81 TGAAGAGAAAACTGAAATGACGACAATAAAAGATGAATTAGGAACTGAAAAAATTAAGAA 140

Qy         168 ACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGA 227
               | | |||||||||||||||||  | || |||   |||| ||||   ||||  || |||||
Db         141 AAATCCTAAACGTGTTGTTGTATTAGAATATAGTTTTGCTGATTATTTAGCAGCATTAGA 200

Qy         228 TGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAG 287
               | | ||||||||||| || || ||||||  ||  |  ||   ||| |  ||  || ||||
Db         201 TATGAAACCTGTTGGTATTGCAGATGATGGCAGCACTAAAAATATAACAAAGTCAGTAAG 260

Qy         288 AGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAAT 347
               |||||| |||||   |||    || || ||| |  |   |||||| ||  | ||||  ||
Db         261 AGATAAGATTGGGGCATATGAATCGGTTGGATCTAGACCGCAACCGAATATGGAAGTGAT 320

Qy         348 CAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAA 407
                |||||| | ||||| ||||| || ||||| |||  ||  ||||| ||    ||  |
Db         321 AAGTAAATTAAAACCGGATTTGATCATTGCAGATGTTAGCAGACATAAGAAAATCAAATC 380

Qy         408 AGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGA 467
               ||| || |  ||||||||||| || ||     |    ||       |||||||||||||
Db         381 AGAATTGAGCAAAATTGCTCCGACAATCATGTTAGTTAGCGGTACGGGAGATTATAATGC 440

Qy         468 AAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAA 527
               ||||||||| || ||||||||| |  | |||||  |||| |||||  |||||||  | ||
Db         441 AAATATTGAAGCATTTAAAACAGTCGCTAAAGCAGTAGGCAAAGAGAAAGAAGGCGAGAA 500

Qy         528 ACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAA 587
                ||  | ||| | || ||||| | | | |  ||   || |||  ||||      |  ||
Db         501 GCGTCTGGAAAAGCATGATAAAATATTAGCGGAGATTAGAAAGAAAATTGAACAGAGTAC 560

Qy         588 AAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAA 647
                    ||   | |||  |    |      | | | ||||| || ||  | ||       |
Db         561 GTTAAAATCTGCATTTGCATTCGGTATCTCAAGAGCAGGTATGTTTATTAATAATGAAGA 620

Qy         648 CTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGT 707
                 | | | | || |||||| ||| | || | ||| || ||   | | | |   | ||
Db         621 TACATTTATGGGACAATTCTTAATTAAAATGGGTATTCAACCTGAAGTCACAAAAGACAA 680

Qy         708 TACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTT 767
                ||||    | |  || |          || |||||  || |  | || | ||||    |
Db         681 AACTACGCATGTTGGTGAACGCAAGGGTGGTCCTTATATATATTTAAATAATGAAGAACT 740

Qy         768 ATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACC 827
                 |  |  | |||||  |   |||| |  ||   ||  ||  |  ||     ||  |
Db         741 TGCCAATATCAATCCAAAAGTTATGATTTTAGCCACTGACGGAAAAACGGACAAAAA--- 797

Qy         828 TTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCA 887
                     ||   || |     |   | | || |||||   |||||| ||||||||| || |
Db         798 ---TAGAACGAAATTCATTGATCCTGCAGTTTGGAAATCATTAAAAGCTGTGAAAGATAA 854

Qy         888 ACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGA 947
                   ||| ||    | |||||  |    |||   | |||| | ||  | ||  |    ||
Db         855 CAAAGTTTATGACGTTGACCGAAATAAGTGGTTGAAATCAAGGGGGATTATCGCAAGTGA 914

Qy         948 AGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAG 985
               |   |||||| ||||  | |   || |  |  | ||||
Db        915 AAGTATGGCAGAAGATTTAGAAAAAATTGCAGAAAAAG 952


RESULT 16
ACA20079
ID   ACA20079 standard; DNA; 960 BP.
XX
AC   ACA20079;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Prokaryotic essential gene #1736.
XX
KW   Antisense; ds; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
XX
OS   Staphylococcus aureus.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,  Zamudio C,  Malone C,  Haselbeck R,  Ohlsen KL,  Zyskind JW;
PI   Wall D,  Trawick JD,  Carr GJ,  Yamamoto R,  Forsyth RA,  Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   P-PSDB; ABU16209.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 14; SEQ ID NO 7949; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential genes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 960 BP; 385 A; 115 C; 210 G; 250 T; 0 U; 0 Other;

  Query Match             26.1%;  Score 263.2;  DB 8;  Length 960;
  Best Local Similarity   55.9%;  Pred. No. 5.4e-39;
  Matches  524;  Conservative    0;  Mismatches  408;  Indels    6;  Gaps    1;

Qy         48 ATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAA 107
              |||  |||| || ||   ||    ||     | |       |||   || |   |  ||
Db         21 ATTAGTTGTGTTTATGCTAATCTTAGTTGTAGCAGTAGCGGGTTGTGGTCAAAAAGATAC 80

Qy        108 AGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAA 167
               ||      ||  ||         ||| ||  | |||  ||| |||   ||| |    ||
Db         81 TGAAGAGAAAACTGAAATGACGACAATAAAAGATGAATTAGGAACTGAAAAAATTAAGAA 140

Qy        168 ACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGA 227
              | | |||||||||||||||||  | || |||   |||| ||||   ||||  || |||||
Db        141 AAATCCTAAACGTGTTGTTGTATTAGAATATAGTTTTGCTGATTATTTAGCAGCATTAGA 200

Qy        228 TGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAG 287
              | | ||||||||||| || || ||||||  ||  |  ||   ||| |  ||  || ||||
Db        201 TATGAAACCTGTTGGTATTGCAGATGATGGCAGCACTAAAAATATAACAAAGTCAGTAAG 260

Qy        288 AGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAAT 347
              |||||| |||||   |||    || || ||| |  |   |||||| ||  | ||||  ||
Db        261 AGATAAGATTGGGGCATATGAATCGGTTGGATCTAGACCGCAACCGAATATGGAAGTGAT 320

Qy        348 CAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAA 407
               |||||| | ||||| ||||| || ||||| |||  ||  ||||| ||    ||  |
Db        321 AAGTAAATTAAAACCGGATTTGATCATTGCAGATGTTAGCAGACATAAGAAAATCAAATC 380

Qy        408 AGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGA 467
              ||| || |  ||||||||||| || ||     |    ||       |||||||||||||
Db        381 AGAATTGAGCAAAATTGCTCCGACAATCATGTTAGTTAGCGGTACGGGAGATTATAATGC 440

Qy        468 AAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAA 527
              ||||||||| || ||||||||| |  | |||||  |||| |||||  |||||||  | ||
Db        441 AAATATTGAAGCATTTAAAACAGTCGCTAAAGCAGTAGGCAAAGAGAAAGAAGGCGAGAA 500

Qy        528 ACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAA 587
               ||  | ||| | || ||||| | | | |  ||   || |||  ||||      |  ||
Db        501 GCGTCTGGAAAAGCATGATAAAATATTAGCGGAGATTAGAAAGAAAATTGAACAGAGTAC 560

Qy        588 AAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAA 647
                   ||   | |||  |    |      | | | ||||| || ||  | ||       |
Db        561 GTTAAAATCTGCATTTGCATTCGGTATCTCAAGAGCAGGTATGTTTATTAATAATGAAGA 620

Qy        648 CTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGT 707
                | | | | || |||||| ||| | || | ||| || ||   | | | |   | ||
Db        621 TACATTTATGGGACAATTCTTAATTAAAATGGGTATTCAACCTGAAGTCACAAAAGACAA 680

Qy        708 TACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTT 767
               ||||    | |  || |          || |||||  || |  | || | ||||    |
Db        681 AACTACGCATGTTGGTGAACGCAAGGGTGGTCCTTATATATATTTAAATAATGAAGAACT 740

Qy        768 ATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACC 827
                |  |  | |||||  |   |||| |  ||   ||  ||  |  ||     ||  |
Db        741 TGCCAATATCAATCCAAAAGTTATGATTTTAGCCACTGACGGAAAAACGGACAAAAA--- 797

Qy        828 TTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCA 887
                    ||   || |     |   | | || |||||   |||||| ||||||||| || |
Db        798 ---TAGAACGAAATTCATTGATCCTGCAGTTTGGAAATCATTAAAAGCTGTGAAAGATAA 854

Qy        888 ACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGA 947
                  ||| ||    | |||||  |    |||   | |||| | ||  | ||  |    ||
Db        855 CAAAGTTTATGACGTTGACCGAAATAAGTGGTTGAAATCAAGGGGGATTATCGCAAGTGA 914

Qy        948 AGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAG 985
              |   |||||| ||||  | |   || |  |  | ||||
Db        915 AAGTATGGCAGAAGATTTAGAAAAAATTGCAGAAAAAG 952
```

```
RESULT 17
ADF30765/c
ID   ADF30765 standard; DNA; 5718 BP.
XX
AC   ADF30765;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Bacillus subtilis strain MB1510 integration region DNA SEQ ID NO:20.
XX
KW   expression library; Gram-positive bacterium host cell;
KW   non-replicating linear integration cassette; gene; Bacillus; ds.
XX
OS   Bacillus subtilis.
XX
PN   WO2003095658-A1.
XX
PD   20-NOV-2003.
XX
PF   07-MAY-2003; 2003WO-DK000301.
XX
PR   07-MAY-2002; 2002DK-00000682.
XX
PA   (NOVO ) NOVOZYMES AS.
XX
PI   Bjornvad ME,  Jorgensen PL,  Hansen PK;
XX
DR   WPI; 2004-012127/01.
XX
PT   Generating an expression library of polynucleotides by introducing the
PT   linear integration cassette into the host cell and selecting or screening
PT   for host cells that produce the polypeptides of interest.
XX
PS   Example 5; SEQ ID NO 20; 55pp; English.
XX
CC   The present invention describes a method for generating an expression
CC   library of polynucleotides integrated by homologous recombination into
CC   the genome of a competent Gram-positive bacterium host cell. The method
CC   comprises: (a) providing a non-replicating linear integration cassette;
CC   and (b) introducing the linear integration cassette into the host cell
CC   and selecting or screening for host cells that produce the polypeptides
CC   of interest. The cassette comprises: (a) a polynucleotide encoding one or
CC   more polypeptides of interest; (b) a 5' flanking polynucleotide segment
CC   upstream of the polynucleotide of (1) and comprising a first homologous
CC   region located in the 3' end of the segment; and (c) a 3' flanking
CC   polynucleotide segment downstream of the polynucleotide of (1) and
CC   comprising a second homologous region located in the 5' end of the
CC   segment. The first and second homologous regions are at least 500, 1000
CC   or 1500 bp, each of which has a sequence identity of at least 80, 85, 90
CC   or 95-100% with a region of the host cell genome. Also described: (1) a
CC   non-replicating linear Gram-positive host cell integration cassette; and
CC   (2) a method of producing a polypeptide of interest. The method is useful
CC   in generating an expression library of polynucleotides integrated by
CC   homologous recombination into the genome of a competent Gram-positive
CC   bacterium host cell. The present sequence represents a Bacillus subtilis
CC   strain MB1510 integration region, which is used in an example from the
CC   present invention.
XX
SQ   Sequence 5718 BP; 1693 A; 1089 C; 1259 G; 1677 T; 0 U; 0 Other;

  Query Match             18.9%;  Score 190.8;  DB 12;  Length 5718;
  Best Local Similarity   55.0%;  Pred. No. 1e-25;
  Matches  488;  Conservative    0;  Mismatches  357;  Indels   43;  Gaps    4;

Qy        115 TCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCT 174
              || || |||      |  ||| |  |||||||||| |  |||| ||||         |||
Db        931 TCTAAGGATTCCAGAGTGATCCATGACGAAGAAGGAAAAACGACAGTAAGCGGCACACCT 872

Qy        175 AAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAA 234
              || || || ||||| ||||||  ||||||  | |||||  |      | | |   |||
Db        871 AAGCGGGTGGTTGTGCTTGAGCTTTCATTCTTGGATGCCGTTCACAATCTCGGCATTACG 812

Qy        235 CCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAA 294
              || || || || || ||||| ||||||||| ||   ||||  || |   |  |
Db        811 CCGGTGGGCATCGCAGATGACAACAAAAAAGATATGATTAAAAAGCTTGTCGGCAGCTCC 752

Qy        295 ATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAA 354
              ||||   | ||||| |||||||| ||||| |   |||| ||  | || |  ||||||
Db        751 ATTG---ATTACACATCTGTAGGCACACGCAGCGAACCCAATCTTGAGGTCATCAGTTCC 695

Qy        355 CTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTA 414
               | || || |||||||| || |||||   | |  | || |||   ||||||||| |  |
Db        694 TTGAAGCCTGATTTAATCATCGCTGACGCTGAGCGCCATAAAAACATTTATAAACAGCTG 635

Qy        415 AATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATT 474
              || ||||| || || ||||||||| | |||||    || |     ||| | ||||  ||
Db        634 AAAAAAATCGCCCCGACGATTGAATTAAAAAGCCGTGAAGCGACATATGACGAAACGATC 575

Qy        475 GATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTA 534
              ||    ||||  || ||| | ||||| |||  ||||||||| ||||| ||| ||    |
Db        574 GACAGCTTTACGACCATTGCTAAAGCATTAAATAAAGAAGATGAAGGAAAAGAAAAGCTT 515

Qy        535 GAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAA 594
              |  || ||| | ||    ||  | ||    |||   ||| |  | |  ||| |||| |
Db        514 GCCGAGCACAAAAAAGTCATCAACGATCTAAAAGCCGAACTTCCGAAAGATGAAAACCGC 455

Qy        595 AAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTAT 654
              ||  |  | |  |  || ||             || |    |||| ||      || |||
Db        454 AACATCGTTCTCGGCGTTGCAAGAGCGGATTCCTTCCAGCTTCATACATCATCATCCTAT 395

Qy        655 GTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAA 714
              |  ||  || || | |     ||||| |||| | | ||  | |     |||
Db        394 GACGGAGAAATCTTTAAAATGCTAGGCTTTACACACGCTGTGAAGTCAGAT--------- 344

Qy        715 GGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAA 774
                                       | ||   | |  | | |  |||    ||     ||
Db        343 ---------------------AACGCCTATCAAGAGGTCAGCCTTGAGCAATTGAGCAAA 305

Qy        775 GTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
               |  |||||||   | |||||||     | ||| |||  |   | |  |         ||
Db        304 ATCGATCCTGATATTTTGTTCATCTCAGCCAACGAAGGCAAAACCATTG---------TA 254

Qy        835 AAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTT 894
               | ||     |||   | ||  | ||||| ||  | || || |||||||||  ||  ||
Db        253 GATGAGTGGAAAACGAACCCGCTCTGGAAAAATCTCAAAGCGGTGAAAAATGGACAAGTC 194

Qy        895 GATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATG 954
               ||  |   |||||||||   ||| ||||||   | ||  | |  |||   |||  ||
Db        193 TATGATGCGGACCGTGACACTTGGACAAGATTCAGAGGCATCAAGTCTAGTGAAACAAGC 134

Qy        955 GCAAAAGAACTTGTTGAATTA-TCTAAGAAAGATAGTAAAAAAGATAA 1001
              || |||||  |  || ||  | | ||  | | |||| | ||| || ||
Db        133 GCCAAAGATGTGCTTAAAAAAGTGTATAATAAATAGGACAAATGAGAA 86


RESULT 18
AAV74666
ID   AAV74666 standard; DNA; 668 BP.
XX
AC   AAV74666;
XX
DT   16-MAR-1999  (first entry)
XX
DE   Staphylococcus aureus contig SEQ ID #355.
XX
KW   Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW   cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW   skin infection; surgical wound infection; scalded skin syndrome;
KW   toxic shock syndrome; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    61. .120
FT                   /*tag= a
FT                   /note= "these bases represent a line of missing text in
FT                   the sequence listing in the specification. They are
FT                   included to maintain the nucleotide numbering given in
FT                   the specification for this DNA sequence"
XX
PN   EP786519-A2.
XX
PD   30-JUL-1997.
XX
PF   07-JAN-1997;   97EP-00100117.
XX
PR   05-JAN-1996;   96US-0009861P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Kunsch CA,  Choi GH,  Barash SC,  Dillon PJ,  Fannon MR,  Rosen CA;
XX
DR   WPI; 1997-374922/35.
XX
PT   Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT   stored on computer readable medium and used in the production of anti-
PT   S.aureus vaccines.
XX
PS   Claim 1; Page 1239-1240; 3271pp; English.
XX
CC   This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC   of the invention. The DNA sequences are recorded on a computer readable
CC   medium, preferably selected from a floppy or hard disk, random access
CC   memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC   the S.aureus DNA sequences allows putative functions to be assigned so
CC   that protein-encoding or regulatory regions of commercial, therapeutic or
CC   industrial importance can be obtained. Specifically, sequences which are
CC   likely to encode antigens have been identified and these polypeptides can
CC   be used in a vaccine composition against S.aureus infection. The
CC   polypeptides can also be used in a kit for the immunodetection of
CC   S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC   including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC   skin and surgical wound infections, scalded skin syndrome, toxic shock
CC   syndrome, etc. Organisms transformed with the DNA sequences can be used
CC   for recombinant production of the polypeptides. The new DNA sequences
CC   (and their fragments) are useful as primers or probes for isolating
CC   homologues of any of the S.aureus DNA sequences contained on the computer
CC   readable medium
XX
SQ   Sequence 668 BP; 216 A; 83 C; 119 G; 190 T; 0 U; 60 Other;

  Query Match             17.4%;  Score 175.4;  DB 2;  Length 668;
  Best Local Similarity   60.9%;  Pred. No. 6e-23;
  Matches  227;  Conservative    0;  Mismatches  146;  Indels    0;  Gaps    0;

Qy          627 TTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAA 686
                |||  | || ||||||| | | || |||||||| |||||  |||   ||||||| |||||
Db            1 TTTATTAGCACATCCAAACTATTCATATGTTGGACAATTTTTAAACGAACTAGGATTTAA 60

Qy          687 AGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTT 746

Db           61 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 120

Qy          747 ACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAA 806
                |||| |  ||||||||  |||| || |  | ||||| ||||||||| |||| |||||| |
Db          121 ACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGATCATTATGACAGA 180

Qy          807 CAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAA 866
                  | || |      |     ||  | | ||  |  || || ||||| |   |||||| ||
Db          181 TCATGCTAAAAAAGATTCTGCTGAATTCAAGAAGTTACAAGAAGATGCAACATGGAAAAA 240

Qy          867 ATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATC 926
                 || || || || |||||| | || || |||||| | ||||||||  | |||||||||||
Db          241 GTTGAATGCAGTTAAAAATAATCGCGTGGATATTGTTGACCGTGATGTTTGGGCAAGATC 300

Qy          927 ACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGA 986
                 ||||| ||||||||||| ||||||||||| |||||||||||||||||||| || |||||
Db          301 TCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTTGTTGAATTATCAAAAAAAGA 360
```

```
Qy          987 TAGTAAAAAAGAT 999
                    ||  ||| |
Db          361 ACAAAAGTAAGGT 373

RESULT 19
AAV76867
ID   AAV76867 standard; DNA; 242 BP.
XX
AC   AAV76867;
XX
DT   16-MAR-1999  (first entry)
XX
DE   Staphylococcus aureus contig SEQ ID #2556.
XX
KW   Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW   cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW   skin infection; surgical wound infection; scalded skin syndrome;
KW   toxic shock syndrome; ds.
XX
OS   Staphylococcus aureus.
XX
PN   EP786519-A2.
XX
PD   30-JUL-1997.
XX
PF   07-JAN-1997;   97EP-00100117.
XX
PR   05-JAN-1996;   96US-0009861P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Kunsch CA,  Choi GH,  Barash SC,  Dillon PJ,  Fannon MR,  Rosen CA;
XX
DR   WPI; 1997-374922/35.
XX
PT   Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT   stored on computer readable medium and used in the production of anti-
PT   S.aureus vaccines.
XX
PS   Claim 1; Page 2272; 3271pp; English.
XX
CC   This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC   of the invention. The DNA sequences are recorded on a computer readable
CC   medium, preferably selected from a floppy or hard disk, random access
CC   memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC   the S.aureus DNA sequences allows putative functions to be assigned so
CC   that protein-encoding or regulatory regions of commercial, therapeutic or
CC   industrial importance can be obtained. Specifically, sequences which are
CC   likely to encode antigens have been identified and these polypeptides can
CC   be used in a vaccine composition against S.aureus infection. The
CC   polypeptides can also be used in a kit for the immunodetection of
CC   S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC   including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC   skin and surgical wound infections, scalded skin syndrome, toxic shock
CC   syndrome, etc. Organisms transformed with the DNA sequences can be used
CC   for recombinant production of the polypeptides. The new DNA sequences
CC   (and their fragments) are useful as primers or probes for isolating
CC   homologues of any of the S.aureus DNA sequences contained on the computer
CC   readable medium
XX
SQ   Sequence 242 BP; 100 A; 31 C; 44 G; 65 T; 0 U; 2 Other;

  Query Match             16.1%;  Score 162.2;  DB 2;  Length 242;
  Best Local Similarity   79.3%;  Pred. No. 1.5e-20;
  Matches  191;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy          208 GATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAAT 267
                ||||| ||||  || ||||| |||||||| ||||| || || ||||||   || |||||
Db            2 GATGCATTAGCAGCATTAGACGTTAAACCAGTTGGTATTGCTGATGATGGTAAGAAAAAA 61

Qy          268 CGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAG 327
```

```
                 ||||| ||||||||| | ||||| ||||||||  | || ||||||||||| ||||||||
Db          62 CGTATCATTAAACCAGTTAGAGAAAAAATTGGGGATTATACTTCTGTAGGTACACGTAAA 121

Qy         328 CAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAAT 387
               || || |||||||||||||| |||||| | ||||| ||||||||||| |||||||  | |
Db         122 CAGCCAAACTTAGAAGAAATTAGTAAATTAAAACCGGATTTAATTATCGCTGATAGCAGT 181

Qy         388 AGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGT 447
               ||||| | ||||||| ||||||| ||||  |||||||| || ||  |   | | || |||
Db         182 AGACATANAGGTATTAATAAAGAATTAANCAAAATTGCACCAACATTATCATTAAAGAGT 241

Qy         448 T 448
               |
Db         242 T 242


RESULT 20
ABN92560
ID   ABN92560 standard; DNA; 183 BP.
XX
AC   ABN92560;
XX
DT   24-JUL-2002  (first entry)
XX
DE   Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:2023.
XX
KW   Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW   antibacterial; gene therapy; gene; ds.
XX
OS   Staphylococcus epidermidis.
XX
PN   US6380370-B1.
XX
PD   30-APR-2002.
XX
PF   13-AUG-1998;   98US-00134001.
XX
PR   14-AUG-1997;   97US-0055779P.
PR   08-NOV-1997;   97US-0064964P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Doucette-Stamm LA,  Bush D;
XX
DR   WPI; 2002-381255/41.
DR   P-PSDB; ABP40015.
XX
PT   Novel isolated nucleic acid encoding a Staphylococcus epidermis
PT   polypeptide, useful for diagnosing and treating bacterial infections.
XX
PS   Disclosure; SEQ ID NO 2023; 267pp; English.
XX
CC   ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC   frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC   given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC   antibacterial activity and can be used in gene therapy. The sequences can
CC   also be used in the diagnosis and treatment of bacterial infections,
CC   particularly S. epidermidis infections. The sequences can be used to
CC   screen for compounds able to interfere with the S. epidermidis life cycle
CC   or inhibit S. epidermidis infection. N.B. The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from the USPTO web site
XX
SQ   Sequence 183 BP; 73 A; 21 C; 39 G; 50 T; 0 U; 0 Other;

  Query Match             15.7%;  Score 158.2;  DB 6;  Length 183;
  Best Local Similarity   98.2%;  Pred. No. 7.7e-20;
  Matches  160;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         846 AAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGA 905
               ||||   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          21 AAAAAGACCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGA 80
```

```
Qy         906 CCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACT 965
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          81 CCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACT 140

Qy         966 TGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
               |||||||||||||||||||||||||||||||||||||||||||
Db         141 TGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 183


RESULT 21
ACA42836
ID   ACA42836 standard; DNA; 897 BP.
XX
AC   ACA42836;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Prokaryotic essential gene #24493.
XX
KW   Antisense; ds; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
XX
OS   Pasteurella multocida.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,   Zamudio C,  Malone C,  Haselbeck R,  Ohlsen KL,  Zyskind JW;
PI   Wall D,   Trawick JD,  Carr GJ,  Yamamoto R,  Forsyth RA,  Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   P-PSDB; ABU38966.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 14; SEQ ID NO 30706; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential genes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 897 BP; 293 A; 176 C; 195 G; 233 T; 0 U; 0 Other;

  Query Match             15.5%;  Score 156.4;  DB 8;  Length 897;
  Best Local Similarity   51.9%;  Pred. No. 1.8e-19;
  Matches  454;  Conservative    0;  Mismatches  381;  Indels   39;  Gaps    3;

Qy        123 TGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGT 182
              ||  |||    ||||  |  || ||||       |   ||  ||     |||||||||||
Db         63 TGCTGTTACCGTCAAAGATCAAAAAGGGGAATTTACCTTAGATAGCGTACCTAAACGTGT 122

Qy        183 TGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGG 242
              |||||   | || ||||| | ||| |||||  | |      | | ||| |  ||||||||
Db        123 TGTTGCCTTAGAATATTCTTATGTGGATGCACTTGCACAAATTGGTGTCAGCCCTGTTGG 182

Qy        243 GATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAA 302
                | || |||||| ||  | ||||  |||||| |  ||  |  || | |||||| |
Db        183 TGTCGCAGATGACAATGATAAAACCCGTATTCTGCAAAAAGTACGCGATAAAGTGCAGCC 242

Qy        303 ATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACC 362
              ||     || || || || |||   ||||| |  |||||||  ||   |  |||||||||
Db        243 ATGGGAATCAGTGGGGACGCGTTCTCAACCGAGTTTAGAAGCGATTTCTGCACTTAAACC 302

Qy        363 AGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAAT 422
              |||||| || ||||| ||| ||||| | ||    |   | ||| ||||  | || |||||
Db        303 AGATTTGATCATTGCCGATGATAATCGCCATTCTGCCGTCTATGAAGAACTCAAAAAAAT 362

Qy        423 TGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTT 482
               || || ||  | |   | ||      | ||   | ||| | |||||  ||||  |
Db        363 CGCGCCGACAGTCGTCTTTAATTCCCGCCATGAGAACTATCAAGAAAACCTTGAAACCGC 422

Qy        483 TAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACA 542
                ||| |||     |    |||||||||  | |||||    ||   ||  | | | ||||
Db        423 ACAAAAAATCGGTGATTTATTAGGTAAATCAAAAGAAATGCAAGCGCGTATTGCAAAACA 482

Qy        543 CGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATT 602
                |  || | ||||  |      | |  |  | |||  |    |||    |||| |   |
Db        483 TAAGCAGGATATTGCGG------ACATCGCCAAAACGTTACCGAAAGGGAAAAAAGCGAT 536

Qy        603 GCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCA 662
                  |  ||  |   | || |    ||   |    ||   ||| | || |||| ||||
Db        537 TATCGGTGTTTCACGTGAAACCCAATTTAATTTATATAATAGCGAATCCTATGCTGGTGG 596

Qy        663 ATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAG 722
               ||  |       ||||||| | ||    ||  |     ||  | ||   ||  |
Db        597 CTTAGTGGAAGTGCTAGGTTATCAAATGCCAAAAGCCCGTGCCGATAACCAACCTAATGC 656

Qy        723 TAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCC 782
              |  |  |  |  || ||                           |  |   ||  || ||
Db        657 TTCGGTTGGTTTAGAACAAG------------------------TGGCGGCAGAAAAGCC 692

Qy        783 TGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACT 842
              ||| |  ||| || ||||     |         |  | | ||   |     || | ||
Db        693 TGATCTGATGATCTTAATCCATTA---------TCGTGATGAGAGTATTGCAAGAAAATG 743

Qy        843 AGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTT 902
               ||||| || |  |||||||| |   |    || || ||||||   |  ||     |
Db        744 GGAAAATGAAGCCTTATGGAAAATTATCCCTGCGGTAAAAAATGGTCAAGTCATCTTAGC 803

Qy        903 AGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGA 962
                |   | | |||||||||||| |||||||| | |   | |  | |||| |||||| |||||
Db        804 TAATGATAATTTATGGGCAAGAGCACGTGGTATTGATGCCGCTGAAGTAATGGCTAAAGA 863

Qy        963 ACTTGTTGAATTATCTAAGAAAGATAGTAAAAAA 996
              | |    || ||    | ||||      ||| ||

Db        864 AGTCCAAGACTTTGTAACGAAATCCGCCAAATAA 897


RESULT 22
ABK74775
ID   ABK74775 standard; DNA; 801 BP.
XX
AC   ABK74775;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Bacillus licheniformis genomic sequence tag (GST) #2066.
XX
KW   Differential gene expression; genomic sequenced tag; GST;
KW   altered culture condition; environmental stress;
KW   physiological provocation; ds.
XX
OS   Bacillus licheniformis.
XX
PN   WO200229113-A2.
XX
PD   11-APR-2002.
XX
PF   05-OCT-2001; 2001WO-US031437.
XX
PR   06-OCT-2000; 2000US-00680598.
PR   27-MAR-2001; 2001US-0279526P.
XX
PA   (NOVO ) NOVOZYMES BIOTECH INC.
PA   (NOVO ) NOVOZYMES AS.
XX
PI   Berka R,  Clausen IG;
XX
DR   WPI; 2002-416684/44.
XX
PT   Monitoring differential expression of several genes in first Bacillus
PT   cell relative to expression of same genes in one or more second Bacillus
PT   cells, by using substrate containing Bacillus genomic sequenced tag
PT   array.
XX
PS   Claim 4; SEQ ID NO 2066; 200pp; English.
XX
CC   The invention describes a method of monitoring differential expression of
CC   genes in a first Bacillus cell relative to expression of the genes in
CC   other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC   isolated from Bacillus cells to a substrate containing array of Bacillus
CC   genomic sequenced tags (GST), examining the array, and determining
CC   relative gene expression by an observed hybridisation reporter signal of
CC   a spot in the array. The method is useful for measuring the expression of
CC   genes in a first Bacillus cell relative to expression of the same genes
CC   in one or more second Bacillus cells. The method is useful for monitoring
CC   global expression of several genes from a Bacillus cell, discovering new
CC   genes, identifying possible functions of unknown open reading frames and
CC   monitoring gene copy number variation and stability. Monitoring changes
CC   in expression of genes may be used to provide a representation of the way
CC   in which Bacillus cells adapt to changes in culture conditions,
CC   environmental stress or other physiological provocation. Extensive follow
CC   -up characterisation is unnecessary, when one spot on an array equals one
CC   gene or one open reading frame, since sequence information is available.
CC   This sequence represents a genomic sequence tag (GST) used in the method
CC   of the invention. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 801 BP; 257 A; 166 C; 204 G; 174 T; 0 U; 0 Other;

  Query Match             14.8%;  Score 149.4;  DB 6;  Length 801;
  Best Local Similarity   51.8%;  Pred. No. 3.4e-18;
  Matches  414;  Conservative    0;  Mismatches  376;  Indels    9;  Gaps    3;

Qy        176 AACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAAC 235
              ||   ||||| |||||||| |    |||| |||||||| |  ||| |   | | |||| |
Db          1 AAAAAGTTGTGGTTCTTGAATTGGGATTTATTGATGCGCTCCTTGATGCGGGTATTAAGC 60
```

```
Qy       236 CTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAA 295
             |||| || || ||||| ||   ||||   ||   |||||   |     | || |  ||||
Db        61 CTGTCGGAATTGCGGACGACGGCAAACCTAAGTTTATTAACGAGAAGGTCAGGGGAAAAA 120

Qy       296 TTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAAC 355
             |   |   || ||||| || ||  | ||    || || | ||| ||| ||||   |    
Db       121 TCAAAGGGTATACTTCAGTCGGTTCGCGCGCCCAGCCAAGCTTTGAAAAAATTGCTTCTT 180

Qy       356 TTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAA 415
             | || || |||||||||||||| |||   |  || ||||  ||  | ||| |  |  |  
Db       181 TAAAGCCCGATTTAATTATTGCCGATTCGAGCAGGCACAGCGGCGTCTATGATAAGCTGT 240

Qy       416 ATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTG 475
               |||||||| || || || |  || ||||  ||  |||  |||||| | || |   |||
Db       241 CGAAAATTGCGCCGACAATCGCGCTCAAAAACTTGAATGCCGATTATCAGGACACGCTTG 300

Qy       476 ATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAG 535
             |||| | |   || ||| | |||||  | || ||||||   | |   |  |||   || |
Db       301 ATGCATCTCTTACGATTGCGAAAGCCGTCGGCAAAGAAAGCGCAATGGAGAAAAAATTGG 360

Qy       536 AAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAA 595
               |||||  |  | ||  | || ||||  ||| |  || |    |     ||| |    |
Db       361 CTGAACATAAACAAAAGCTGGATGAATTGAAACAGAAATTCGGCAGCCGGAAACAGAGCA 420

Qy       596 AGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATG 655
                |  |||  |      |   | ||  |       |     ||  |   |||| ||   
Db       421 TTCTTCTGCTCGGGAACACAAATGAAGAAATCACCGTGCGCGATGAA---AACTTTTTCA 477

Qy       656 TTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAG 715
                 |||  | || |   || | || | || | | |   |  | ||    |     ||||
Db       478 CGTCTCAGCTTCTGACGAAAATCGGCTATACATACGGTGTCGGAGACAGCG---GCAAAG 534

Qy       716 GTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAG 775
             |        |  |   | ||       |  |  ||||||     |||    |   | |  
Db       535 GCGATGCCGAAAACGGTGAATCCGTCAATATTAAAATGACGCTCGAACAGCTGCTTGAGA 594

Qy       776 TGAATCCTGAGCGTATGTTCATAATGAC---AAACAAAGCAAGTTCTAACGAACCTTCAC 832
                |||| ||   |||  || | |||||   | | ||||  |      | |||       
Db       595 AAGATCCCGACGTTATTGTCCTGATGACGGGAGAAAAAGATAAAGTCGATGAAGACGGGA 654

Qy       833 TAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTG 892
              || |    | ||||||||||||| | ||||| ||  | | ||| || |||||       |
Db       655 AAAGACCGATCGAAAAAGATCCTCTTTGGAAAAAGCTCAGCGCAGTCAAAAACGGCAAGG 714

Qy       893 TTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAA 952
             || ||      ||| |   |   ||| |      |||  |  |         |||| ||  
Db       715 TTTATGAAGCCGACAGATTCGCCTGGTCGCTCCGACGCAGCATTGACGGAGCAGATGAGC 774

Qy       953 TGGCAAAAGAACTTGTTGA 971
             ||    | ||| | | | |
Db       775 TGATGGATGAAATCGATCA 793

RESULT 23
ACA21545
ID   ACA21545 standard; DNA; 972 BP.
XX
AC   ACA21545;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Prokaryotic essential gene #3202.
XX
KW   Antisense; ds; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
XX
OS   Bacillus anthracis.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
```

```
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,   Zamudio C,   Malone C,   Haselbeck R,   Ohlsen KL,   Zyskind JW;
PI   Wall D,   Trawick JD,   Carr GJ,   Yamamoto R,   Forsyth RA,   Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   P-PSDB; ABU17675.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 14; SEQ ID NO 9415; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential genes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 972 BP; 405 A; 147 C; 183 G; 237 T; 0 U; 0 Other;

  Query Match             14.2%;  Score 143.2;  DB 8;  Length 972;
  Best Local Similarity   55.0%;  Pred. No. 4.7e-17;
  Matches  308;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    1;

Qy        25 AAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAAT 84
             |||||| | ||| |  |     || | ||| | ||   ||        | |     ||| 
Db         7 AAAATTCTCAGTATTTTCATAGTAGTTTTTCTATTCGCTGTTGGATGCGGACAGCAAAAA 66

Qy        85 AATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAA 144
              |      ||   |  |||     ||      |||  | |   |   ||| || ||||  
Db        67 GAGGAGAAAAAAGAAACAAAAGCGGACAATAAAAATCAAGCTATAACAATTAAACACGCT 126

Qy       145 GAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTT 204
             |||||    |||||  ||  |||||   | |||   || ||||| ||||| |    || |
Db       127 GAAGGGGAAACGAAGTTAGATAAACCAGCGAAAAAAGTAGTTGTACTTGAATGGGTATAT 186

Qy       205 GTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAA 264
```

```
                 || |  ||| | ||  | | |||| | || || ||||| || ||   ||| ||  |
Db        187 TCAGAAGACTTATTAGCACTTGGTGTTCAGCCAGTAGGGATGGCAGACATTAAGAATTAT 246

Qy        265 AATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGT 324
              |||       | ||  ||  || | |    | |  | |         ||| || ||||||
Db        247 AATAAATGGGTAAATACAAAAACAAAACCGAGTAAAGATGTTGTAGATGTCGGGACACGT 306

Qy        325 AAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAAT 384
               | ||||| |||||||||||||| ||     | |||||||||||||||||  | | |
Db        307 CAACAACCAAACTTAGAAGAAATTAGCCGTTTAAAACCAGATTTAATTATCACAGCTTCA 366

Qy        385 AATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGAT---------T 435
                  |    ||||  ||| | || || ||| |  ||||||| || ||  |         |
Db        367 TTCCGTGGTAAAGCAATTAAAAATGAATTAGAACAAATTGCACCAACAGTTATGTTTGAT 426

Qy        436 GAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCA 495
                |   | |||     | |   | | |  |||||  |  ||  | |||||||  |||| ||
Db        427 CCATCAACAAGCAATAACGATCACTTTGCTGAAATGACAGAAACATTTAAACAAATTGCA 486

Qy        496 AAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATT 555
              |||||  | || |||||||||||||||||||||   ||||  ||    |||||   | |
Db        487 AAAGCAGTTGGAAAAGAAGAAGAAGGTAAAAAAGTATTAGCTGATATGGATAAAGCATTC 546

Qy        556 GAAGAATATAAAAAAGAAAT 575
              |  ||    |||  | ||||
Db        547 GCTGATGCAAAAGCAAAAAT 566

RESULT 24
AAS49530/c
ID   AAS49530 standard; DNA; 321 BP.
XX
AC   AAS49530;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus cellular proliferation inhibitory sequence #754.
XX
KW   Antisense; ss; prokaryotic cellular proliferation; antibiotic;
KW   antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US009180.
XX
PR   21-MAR-2000; 2000US-0191078P.
PR   23-MAY-2000; 2000US-0206848P.
PR   26-MAY-2000; 2000US-0207727P.
PR   23-OCT-2000; 2000US-0242578P.
PR   27-NOV-2000; 2000US-0253625P.
PR   22-DEC-2000; 2000US-0257931P.
PR   16-FEB-2001; 2001US-0269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids.
XX
PS   Claim 1; SEQ ID NO 2107; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the genes,
CC   their use in the discovery of novel antibiotics, the essential genes
CC   themselves and the encoded proteins. The prokaryotes used are Escherichia
CC   coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC   Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC   useful for the identification of potential new targets for antibiotic
CC   development. The antisense nucleic acids can also be used to identify
CC   proteins used in proliferation, to express these proteins, and to obtain
CC   antibodies capable of binding to the expressed proteins. The proteins can
CC   be used to screen compounds in rational drug discovery programmes. The
CC   antisense nucleic acid sequence is also useful to screen for homologous
CC   nucleic acids which are required for cell proliferation in a wide variety
CC   of organisms. The present sequence is an antisense oligonucleotide of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 321 BP; 83 A; 73 C; 36 G; 129 T; 0 U; 0 Other;

  Query Match             12.7%;  Score 128;  DB 4;  Length 321;
  Best Local Similarity   67.8%;  Pred. No. 2.7e-14;
  Matches  179;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy        132 AATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCT 191
              ||| ||  | |||  ||| |||   ||| |    ||| | |||||||||||||||||  |
Db        271 AATAAAAGATGAATTAGGAACTGAAAAAATTAAGAAAAATCCTAAACGTGTTGTTGTATT 212

Qy        192 TGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGA 251
               || |||   |||| ||||   ||||  || |||||| | ||||||||||| || || ||
Db        211 AGAATATAGTTTTGCTGATTATTTAGCAGCATTAGATATGAAACCTGTTGGTATTGCAGA 152

Qy        252 TGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTC 311
              ||||  ||  |  ||   ||| |  ||  || |||||||||| |||||   |||    ||
Db        151 TGATGGCAGCACTAAAAATATAACAAAGTCAGTAAGAGATAAGATTGGGGCATATGAATC 92

Qy        312 TGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAAT 371
               || ||| |  |   |||||| ||  | ||||  || |||||| | ||||| ||||| ||
Db         91 GGTTGGATCTAGACCGCAACCGAATATGGAAGTGATAAGTAAATTAAAACCGGATTTGAT 32

Qy        372 TATTGCTGATAATAATAGACACAA 395
               ||||| |||  ||  ||||| ||
Db         31 CATTGCAGATGTTAGCAGACATAA 8

RESULT 25
ACA16735/c
ID   ACA16735 standard; DNA; 321 BP.
XX
AC   ACA16735;
XX
DT   27-OCT-2003  (revised)
DT   19-JUN-2003  (first entry)
XX
DE   Prokaryotic essential gene antisense oligonucleotide #4605.
XX
KW   Antisense; ss; prokaryotic essential gene; cell proliferation;
KW   drug design.
XX
OS   Archaea.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,  Zamudio C,  Malone C,  Haselbeck R,  Ohlsen KL,  Zyskind JW;
PI   Wall D,  Trawick JD,  Carr GJ,  Yamamoto R,  Forsyth RA,  Xu HH;
```

```
XX
DR   WPI; 2003-029926/02.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 1; SEQ ID NO 4605; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the 6213
CC   antisense sequences of the invention. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences. (Updated on 27-OCT-2003 to
CC   standardise OS field)
XX
SQ   Sequence 321 BP; 83 A; 73 C; 36 G; 129 T; 0 U; 0 Other;

  Query Match             12.7%;  Score 128;  DB 8;  Length 321;
  Best Local Similarity   67.8%;  Pred. No. 2.7e-14;
  Matches ·179;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy        132 AATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCT 191
              ||| ||  | |||  ||| |||   ||| |    ||| | |||||||||||||||||  |
Db        271 AATAAAAGATGAATTAGGAACTGAAAAAATTAAGAAAAATCCTAAACGTGTTGTTGTATT 212

Qy        192 TGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGA 251
               || |||   |||| ||||   ||||  || |||||| | ||||||||||| || || ||
Db        211 AGAATATAGTTTTGCTGATTATTTAGCAGCATTAGATATGAAACCTGTTGGTATTGCAGA 152

Qy        252 TGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTC 311
              ||||  ||  |  ||   ||| |  ||  || |||||||||| |||||   |||    ||
Db        151 TGATGGCAGCACTAAAAATATAACAAAGTCAGTAAGAGATAAGATTGGGGCATATGAATC 92

Qy        312 TGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAAT 371
               || ||| |  |   |||||| ||  | ||||  || |||||| | ||||| ||||| ||
Db         91 GGTTGGATCTAGACCGCAACCGAATATGGAAGTGATAAGTAAATTAAAACCGGATTTGAT 32

Qy        372 TATTGCTGATAATAATAGACACAA 395
               ||||| |||  ||  ||||| ||
Db         31 CATTGCAGATGTTAGCAGACATAA 8


RESULT 26
AAV74915
ID   AAV74915 standard; DNA; 2115 BP.
XX
```

```
AC   AAV74915;
XX
DT   16-MAR-1999  (first entry)
XX
DE   Staphylococcus aureus contig SEQ ID #604.
XX
KW   Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW   cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW   skin infection; surgical wound infection; scalded skin syndrome;
KW   toxic shock syndrome; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    541. .600
FT                   /*tag=  a
FT                   /note= "these bases represent a line of missing text in
FT                   the sequence listing in the specification. They are
FT                   included to maintain the nucleotide numbering given in
FT                   the specification for this DNA sequence"
XX
PN   EP786519-A2.
XX
PD   30-JUL-1997.
XX
PF   07-JAN-1997;   97EP-00100117.
XX
PR   05-JAN-1996;   96US-0009861P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Kunsch CA,  Choi GH,  Barash SC,  Dillon PJ,  Fannon MR,  Rosen CA;
XX
DR   WPI; 1997-374922/35.
XX
PT   Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT   stored on computer readable medium and used in the production of anti-
PT   S.aureus vaccines.
XX
PS   Claim 1; Page 1535-1536; 3271pp; English.
XX
CC   This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC   of the invention. The DNA sequences are recorded on a computer readable
CC   medium, preferably selected from a floppy or hard disk, random access
CC   memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC   the S.aureus DNA sequences allows putative functions to be assigned so
CC   that protein-encoding or regulatory regions of commercial, therapeutic or
CC   industrial importance can be obtained. Specifically, sequences which are
CC   likely to encode antigens have been identified and these polypeptides can
CC   be used in a vaccine composition against S.aureus infection. The
CC   polypeptides can also be used in a kit for the immunodetection of
CC   S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC   including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC   skin and surgical wound infections, scalded skin syndrome, toxic shock
CC   syndrome, etc. Organisms transformed with the DNA sequences can be used
CC   for recombinant production of the polypeptides. The new DNA sequences
CC   (and their fragments) are useful as primers or probes for isolating
CC   homologues of any of the S.aureus DNA sequences contained on the computer
CC   readable medium
XX
SQ   Sequence 2115 BP; 787 A; 279 C; 363 G; 622 T; 0 U; 64 Other;

  Query Match             12.4%;  Score 125.4;  DB 2;  Length 2115;
  Best Local Similarity   53.2%;  Pred. No. 8.9e-14;
  Matches  289;  Conservative    1;  Mismatches  247;  Indels    6;  Gaps    1;

Qy        458 ATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAG 517
              ||||||||| ||||||||| || ||||||||| |  | |||||  |||| |||||  |||
Db          1 ATTATAATGCAAATATTGAAGCATTTAAAACAGTCGCTAAAGCAGTAGGCAAAGAGAAAG 60

Qy        518 AAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAA 577
              ||||  | || ||  | ||| | || ||||| | | | |  ||   || |||  ||||
Db         61 AAGGCGAGAAGCGTCTGGAAAAGCATGATAAAATATTAGCGGAGATTAGAAAGAAAATTG 120
```

```
Qy          578 CTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTC 637
                    |  ||      ||   | |||  |    |      | | | ||||| || ||  |
Db          121 AACAGAGTACGTTAAAATCTGCATTTGCATTCGGTATCTCAAGAGCAGGTATGTTTATTA 180

Qy          638 ATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAA 697
                ||       |  | | | | || |||||| ||| | || | ||| || ||   | | | |
Db          181 ATAATGAAGATACATTTATGGGACAATTCTTAATTAAAATGGGTATTCAACCTGAAGTCA 240

Qy          698 GTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACA 757
                   | :|    ||||    | |  || |          || |||||  || |  | || |
Db          241 MAAAARAMAAAACTACGCATGTTGGTGAACGCAAGGGTGGTCCTTATATATATTTAAATA 300

Qy          758 CTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTT 817
                 ||||    |  |  |  | |||||  |   |||| |  ||   ||  ||  |  ||
Db          301 ATGAAGAACTTGCCAATATCAATCCAAAAGTTATGATTTTAGCCACTGACGGAAAAACGG 360

Qy          818 CTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTG 877
                  ||  |         ||   || |     |   | | || |||||   |||||| ||||
Db          361 ACAAAAA------TAGAACGAAATTCATTGATCCTGCAGTTTGGAAATCATTAAAAGCTG 414

Qy          878 TGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAA 937
                ||||| || |    ||| ||    | |||||  |    |||   | |||| | ||  | |
Db          415 TGAAAGATAACAAAGTTTATGACGTTGACCGAAATAAGTGGTTGAAATCAAGGGGGATTA 474

Qy          938 TTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAG 997
                |  |    |||   |||||| ||||  | |   || |  |  | ||||  |   |||||
Db          475 TCGCAAGTGAAAGTATGGCAGAAGATTTAGAAAAAATTGCAGAAAAAGCAAAATAAAAAT 534

Qy          998 ATA 1000
                | |
Db          535 ACA 537

RESULT 27
ABT15032
ID   ABT15032 standard; DNA; 990 BP.
XX
AC   ABT15032;
XX
DT   06-MAR-2003  (first entry)
XX
DE   Pathogen specific antigen related staphylococcal DNA SEQ ID No 318.
XX
KW   Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;
KW   hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
KW   auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
KW   autoimmune disease; HIV; hepatitis; gene; ds.
XX
OS   Staphylococcus sp.
XX
PN   WO200259148-A2.
XX
PD   01-AUG-2002.
XX
PF   21-JAN-2002; 2002WO-EP000546.
XX
PR   26-JAN-2001; 2001AT-00000130.
XX
PA   (CIST-) CISTEM BIOTECHNOLOGIES GMBH.
XX
PI   Meinke A,  Nagy E,  Von Ahsen U,  Klade C,  Henics T,  Zauner W;
PI   Minh DB,  Vytvytska O,  Etz H,  Dryla A,  Weichhart T,  Hafner M;
PI   Tempelmaier B;
XX
DR   WPI; 2003-075410/07.
XX
PT   Identifying, isolating and producing hyperimmune serum-reactive antigens
PT   from a pathogen, for preparing vaccine or medicament for treating or
PT   preventing e.g. staphylococcal infections, comprises providing antibody
PT   preparation.
XX
PS   Example 7; Page 210; 252pp; English.
XX
CC   The invention relates to a novel method for identifying, isolating and
CC   producing hyperimmune serum-reactive antigens from a pathogen, tumour,
CC   allergen, a tissue or host prone to auto-immunity, where the antigens are
CC   used in a vaccine, comprises providing antibody preparation from a plasma
CC   pool of a type of animal, or individual sera with antibodies against the
CC   specific pathogen, tumour, allergen, tissue or host prone to auto-
CC   immunity. The hyperimmune serum-reactive antigens comprising any of the
CC   62 sequences of 53-2261 amino acids fully defined in the specification,
CC   or their hyperimmune fragments are useful for the manufacture of a
CC   pharmaceutical preparation, particularly a vaccine against staphylococcal
CC   infections or colonisation against S. aureus or S. epidermidis. The
CC   preparation of antibodies is useful for the manufacture of a medicament
CC   for treating or preventing staphylococcal infections or colonisation
CC   against S. aureus or S. epidermidis. The antibody preparations may also
CC   be used for diagnostic and imaging purposes. Other conditions that can be
CC   treated include cancer, autoimmune diseases or infections caused by viral
CC   (e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This
CC   polynucleotide sequence represents staphylococcal DNA relating to the
CC   method for identifying and producing pathogen specific antigens of the
CC   invention
XX
SQ   Sequence 990 BP; 398 A; 141 C; 179 G; 272 T; 0 U; 0 Other;

  Query Match             12.3%;  Score 124;  DB 8;  Length 990;
  Best Local Similarity   50.7%;  Pred. No. 1.5e-13;
  Matches  412;  Conservative    0;  Mismatches  380;  Indels   20;  Gaps    4;

Qy           27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
                ||||  ||   |  |||      |   | |  || |||     |||| |||| |    ||
Db           12 AATTAAAATGCTTGTTGTTACGCTTGCTTTCCTACTTGTTTTAGCAGGATGTAGTGGGAA 71

Qy           87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
                |      |   | |||  | |  |   |   || ||   |  |  ||| || || | |
Db           72 TTCAAATAAACAATCATCTGATAACAAAGATAAGGAAACAACTTCAATTAAACATGCAAT 131

Qy          147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
                 ||||| ||  || |        | || || |||||||||   ||  |  |
Db          132 GGGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCAC 191

Qy          207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
                ||| |     ||  |||||| |||||||||||| ||    |  ||       | | ||||
Db          192 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAA 251

Qy          267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
                 |  |   |  ||    ||| | || |  |   | |    |    ||||||   |
Db          252 ACCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACC 311

Qy          327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                   |||||||||||| ||||||  |||| | ||||| || |||||| | ||    ||
Db          312 TGCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGT 371

Qy          387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
                |||| |  ||    ||||  |  | |||  |||||| || ||             || ||
Db          372 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAG 418

Qy          447 TTTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAG 505
                |||| |  || |   |  | ||| | |  ||| |   |||   |||    ||||||||||
Db          419 TTTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAG 478

Qy          506 GTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATA 565
                | |||||| |||||| | || |     |  || |  ||||| | ||| | |  | ||
Db          479 GGAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCC 538

Qy          566 AAAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAAT 622
                |||||||   || | | | | ||||| ||  |  |  ||||   ||   || || ||| |
Db          539 AAAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACT 598

Qy          623 CAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
                   ||     || |||| |||| |  | |  || |||  |              ||  |
Db          599 TCCGT---GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 655

Qy          683 TTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTT 742
```

```
               |  || |     | || || | |      ||     |  | |  ||  |||||    |
Db        656 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 715

Qy        743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGA 802
               |  ||  |     |  ||||   |  |   | ||||  |||| | ||| ||  || | |
Db        716 TCCAACTTACATCTAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAA 775

Qy        803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
               |    |  ||| | | || ||  || || ||
Db        776 AATCAGATCCAAATGCGAAAGATGCTGCATTA 807

RESULT 28
AAZ22850
ID   AAZ22850 standard; DNA; 999 BP.
XX
AC   AAZ22850;
XX
DT   20-DEC-1999  (first entry)
XX
DE   Staphylococcus aureus cbrA DNA.
XX
KW   Infection; detection; diagnosis; screening; antibiotic; resistance;
KW   methicillin; MRSA; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   CDS             7. .999
FT                   /*tag= a
FT                   /product= "S. aureus cbrA protein"
XX
PN   WO9947662-A1.
XX
PD   23-SEP-1999.
XX
PF   18-MAR-1999;   99WO-US006199.
XX
PR   20-MAR-1998;   98US-0078682P.
PR   01-APR-1998;   98US-0080296P.
PR   07-MAY-1998;   98US-0084674P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Simpson AJG,  Choi GH;
XX
DR   WPI; 1999-580304/49.
DR   P-PSDB; AAY42331.
XX
PT   Staphylococcus aureus genes and polypeptides, vectors and methods of
PT   recombinant production.
XX
PS   Claim 1; Page 24; 111pp; English.
XX
CC   This sequence represents Staphylococcus aureus cbrA DNA. The cbrA protein
CC   is predicted to have a molecular weight of about 36.8 kD and, along with
CC   cbrB (AAY42332) and cbrC (AAY42333), is thought to be involved in iron
CC   regulation, based on amino acid sequence homology with known iron
CC   regulator proteins. S. aureus is a ubiquitous pathogen which causes
CC   infections in burns, cellulitis, eyelid infections, food poisoning, joint
CC   infections, neonatal conjunctivitis, osteomyelitis, skin infections,
CC   surgical wound infection, scalded skin syndrome and toxic shock syndrome.
CC   S. aureus is increasingly becoming resistant to known antibiotics, with
CC   methicillin-resistant strains generally being multiply drug resistant.
CC   Methicillin-resistant S. aureus (MRSA) poses serious infection control
CC   problems, with many strains being multiresistant against virtually all
CC   antibiotics with the exception of the vancomycin-type glycopeptide
CC   antibiotics. The protein may be useful to screen potential antagonists
CC   which could be used as antibiotics and it may be used as a vaccine to
CC   prevent or attenuate an infection caused by a member of the
CC   Staphylococcus genus. The protein, or antibodies against it can be used
CC   in immunoassays to detect Staphylococcus in a biological sample. Probes
CC   and primers derived from the nucleic acid sequences may also be used to
CC   detect Staphylococcus nucleic acids in a biological sample
XX
SQ   Sequence 999 BP; 402 A; 142 C; 180 G; 275 T; 0 U; 0 Other;

  Query Match             12.3%;  Score 124;  DB 2;  Length 999;
  Best Local Similarity   50.7%;  Pred. No. 1.5e-13;
  Matches  412;  Conservative    0;  Mismatches  380;  Indels   20;  Gaps    4;

Qy         27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
              ||||  ||   |  |||      |   | |  || |||     |||| |||| |    ||
Db         18 AATTAAAATGCTTGTTGTTACGCTTGCTTTCCTACTTGTTTTAGCAGGATGTAGTGGGAA 77

Qy         87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
              |      |   | |||  | |  |   |   || ||   |  |  ||| || || | |
Db         78 TTCAAATAAACAATCATCTGATAACAAAGATAAGGAAACAACTTCAATTAAACATGCAAT 137

Qy        147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
               ||||| ||  ||  |       | || || |||||||||   ||  |  |
Db        138 GGGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCAC 197

Qy        207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
              ||| |     ||  |||||| ||||||||||||| ||    |  ||       | | ||||
Db        198 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAA 257

Qy        267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
               |  |   |  ||    ||| | || |  |   | |    |    ||||||   |
Db        258 ACCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACC 317

Qy        327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                 |||||||||||| ||||||  |||| | ||||| || |||||| | ||    ||
Db        318 TGCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGT 377

Qy        387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
              |||| |  ||    ||||  |  | |||  |||||| || ||             || ||
Db        378 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAG 424

Qy        447 TTTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAG 505
              |||| |  || |   |  | ||| | |  ||| |   |||   |||    ||||||||||
Db        425 TTTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAG 484

Qy        506 GTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATA 565
              | |||||| |||||| | || |     |  || |  ||||| | ||| | |  | ||
Db        485 GGAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCC 544

Qy        566 AAAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAAT 622
              |||||||   || | | | | ||||| ||  |  |  |||   ||   || || ||| |
Db        545 AAAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACT 604

Qy        623 CAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
                 ||     || |||| |||| |  | |  || |||  |              ||  |
Db        605 TCCGT---GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 661

Qy        683 TTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTT 742
               |  || |     | || || | |      ||     |  | |  ||  |||||    |
Db        662 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 721

Qy        743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGA 802
               |  ||  |     |  ||||   |  |   | ||||  |||| | ||| ||  || | |
Db        722 TCCAACTTACATCTAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAA 781

Qy        803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
               |    |  ||| | | || ||  || || ||
Db        782 AATCAGATCCAAATGCGAAAGATGCTGCATTA 813

RESULT 29
AAZ19889
ID   AAZ19889 standard; DNA; 999 BP.
XX
AC   AAZ19889;
XX
DT   06-DEC-1999  (first entry)
```

```
XX
DE   Staphylococcus aureus iron regulation gene cbrA.
XX
KW   CbrA gene; infection; therapy; diagnosis; vaccine; antibiotic;
KW   iron regulation; ss.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   CDS             7. .999
FT                   /*tag=  a
XX
PN   WO9947639-A2.
XX
PD   23-SEP-1999.
XX
PF   19-MAR-1999;   99WO-US005976.
XX
PR   20-MAR-1998;   98US-0078682P.
PR   01-APR-1998;   98US-0080296P.
PR   07-MAY-1998;   98US-0084674P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Simpson AJG,  Choi GH;
XX
DR   WPI; 1999-562101/47.
DR   P-PSDB; AAY31824.
XX
PT   New isolated Staphylococcus nucleic acid molecules, used to develop
PT   products for the diagnosis, prevention and treatment of Staphylococcal
PT   infections.
XX
PS   Claim 1; Page 24; 102pp; English.
XX
CC   This is the nucleotide sequence of Staphylococcus aureus strain ISP3
CC   (ATCC 202108) genomic DNA including the novel cbrA gene that codes for a
CC   330-amino acid protein (see AAY31824) of predicted mol.wt. 36.8 kDa. The
CC   sequence was obtained from overlapping clones BTACA44 and BTAGJ54, which
CC   span a single operon containing the cbrA, cbrB and cbrC genes (see
CC   AAZ19889-91). CbrA shows sequence homology to known genes involved in
CC   iron regulation. The invention provides 11 novel genes (see AAZ19882-92)
CC   of S. aureus and the polypeptides they encode (see AAY31817-27). Also
CC   provided are vectors, host cells, antibodies and hybridomas. The
CC   invention further relates to screening methods for identifying agonists
CC   and antagonists of S. aureus polypeptide activity, and to diagnostic
CC   methods for detecting Staphylococcus nucleic acids, polypeptides and
CC   antibodies in a biological sample. Antagonists of cbrA may be useful as
CC   antibiotics to treat infections of S. aureus and other Staphylcoccus spp.
CC   Also provided are novel vaccines for the prevention or attenuation of
CC   infection by Staphylococcus. The isolated nucleic acid molecule is also
CC   useful for generating probes and primers, and in the recombinant
CC   production of cbrA protein
XX
SQ   Sequence 999 BP; 402 A; 142 C; 180 G; 275 T; 0 U; 0 Other;

  Query Match             12.3%;  Score 124;  DB 2;  Length 999;
  Best Local Similarity   50.7%;  Pred. No. 1.5e-13;
  Matches  412;  Conservative    0;  Mismatches  380;  Indels   20;  Gaps    4;

Qy         27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
              ||||  ||   |  |||      |   | |  || |||     |||| |||| |    ||
Db         18 AATTAAAATGCTTGTTGTTACGCTTGCTTTCCTACTTGTTTTAGCAGGATGTAGTGGGAA 77

Qy         87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
              |      |   | |||  | |  |   |   || ||   |  |  ||| || || | |
Db         78 TTCAAATAAACAATCATCTGATAACAAAGATAAGGAAACAACTTCAATTAAACATGCAAT 137

Qy        147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
               ||||| ||  || |        | || || |||||||||   ||  |  |
Db        138 GGGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCAC 197
```

```
Qy        207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
              ||| |     ||  |||||| |||||||||||| ||    |  ||       | | ||||
Db        198 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAA 257

Qy        267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
               |  | | |  ||    ||| | || |  |   | |    |    ||||||   |
Db        258 ACCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACC 317

Qy        327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                 |||||||||||| ||||||  |||| | ||||| || |||||| | ||    ||
Db        318 TGCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGT 377

Qy        387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
              |||| |  ||    ||||  |  | |||  |||||| || ||             || ||
Db        378 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAG 424

Qy        447 TTTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAG 505
              |||| |  || |   |  | ||| | |  ||| |   |||   |||    ||||||||||
Db        425 TTTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAG 484

Qy        506 GTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATA 565
              | |||||| |||||| | || |     |  || |  ||||| | ||| | |  | ||
Db        485 GGAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCC 544

Qy        566 AAAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAAT 622
              |||||||   || | | | | ||||| ||  |  |  ||||   ||   || || ||| |
Db        545 AAAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACT 604

Qy        623 CAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
                 ||     || |||| |||| |  | |  || |||  |              ||  |
Db        605 TCCGT---GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 661

Qy        683 TTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTT 742
               |  || |     | || || | |      ||     |  | |  ||  |||||    |
Db        662 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 721

Qy        743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGA 802
               |  ||  |     |  ||||   |  |   | ||||  |||| | ||| ||  || | |
Db        722 TCCAACTTACATCTAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAA 781

Qy        803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
               |    |  ||| | | || ||  || || ||
Db        782 AATCAGATCCAAATGCGAAAGATGCTGCATTA 813


RESULT 30
ADF43555
ID   ADF43555 standard; DNA; 999 BP.
XX
AC   ADF43555;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Staphylococcus aureus cbtrA DNA.
XX
KW   antibacterial; immunostimulant; vaccine; vaccine;
KW   Staphylococcus aureus infection; infection prevention;
KW   infection attenuation; gene; ds; cbrA.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   CDS             7. .999
FT                   /*tag=  a
FT                   /product= "cbrA"
XX
PN   US2003153733-A1.
XX
PD   14-AUG-2003.
XX
PF   24-OCT-2002; 2002US-00278946.
XX
PR   20-MAR-1998;   98US-0078862P.
PR   01-APR-1998;   98US-0080296P.
PR   07-MAY-1998;   98US-0084674P.
PR   18-MAR-1999;   99WO-US006199.
PR   15-JAN-2002; 2002US-00830217.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Simpson AJG,  Choi GH;
XX
DR   WPI; 2003-897737/82.
DR   P-PSDB; ADF43556.
XX
PT   New Staphylococcus aureus gene, useful for preparing a vaccine against
PT   infection caused by Staphylococcus aureus.
XX
PS   Claim 1; SEQ ID NO 15; 61pp; English.
XX
CC   The invention describes a new isolated nucleic acid molecule comprising:
CC   a sequence that is 95% identical to a fully defined sequence having 586-
CC   2226 bp, or its complement; or a sequence encoding a polypeptide having a
CC   fully defined sequence comprising 136-691 amino acids, or its complement.
CC   The gene is Staphylococcus aureus gene. The nucleic acid is useful for
CC   preparing a vaccine against infection caused by Staphylococcus aureus.
CC   The methods are useful for preventing or attenuating an infection caused
CC   by a Staphylococcus, detecting Staphylococcus nucleic acids in a
CC   biological sample obtained from an animal, and detecting Staphylococcus
CC   antibodies in a biological sample obtained from an animal. This sequence
CC   encodes Staphylococcus aureus cbrA protein.
XX
SQ   Sequence 999 BP; 402 A; 142 C; 180 G; 275 T; 0 U; 0 Other;

  Query Match             12.3%;  Score 124;  DB 10;  Length 999;
  Best Local Similarity   50.7%;  Pred. No. 1.5e-13;
  Matches  412;  Conservative    0;  Mismatches  380;  Indels   20;  Gaps    4;

Qy         27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
              ||||  ||   |  |||      |   | |  || |||     |||| |||| |    ||
Db         18 AATTAAAATGCTTGTTGTTACGCTTGCTTTCCTACTTGTTTTAGCAGGATGTAGTGGGAA 77

Qy         87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
              |      |   | |||  | |  |   |   || ||   |  |  ||| || || | |
Db         78 TTCAAATAAACAATCATCTGATAACAAAGATAAGGAAACAACTTCAATTAAACATGCAAT 137

Qy        147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
               ||||| ||  || |        | || || |||||||||   ||  |  |
Db        138 GGGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCAC 197

Qy        207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
              ||| |     ||  |||||| |||||||||||| ||    |  ||       | | ||||
Db        198 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAA 257

Qy        267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
               |  |   |  ||    ||| | || |  |   | |    |    ||||||   |
Db        258 ACCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACC 317

Qy        327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                 |||||||||||| ||||||  |||| | ||||| || |||||| | ||    ||
Db        318 TGCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGT 377

Qy        387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
              |||| |  ||    ||||  |  | |||  |||||| || ||             || ||
Db        378 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAG 424

Qy        447 TTTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAG 505
              |||| |  || |   |  | ||| | |  ||| |   |||   |||    ||||||||||
Db        425 TTTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAG 484

Qy        506 GTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATA 565
              | |||||| |||||| | || |     |  || |  ||||| | ||| | |  | ||
Db        485 GGAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCC 544

Qy        566 AAAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAAT 622
              |||||||   || | | | | ||||| ||  |  |  ||||   ||   || || ||| |
Db        545 AAAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACT 604

Qy        623 CAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
                 ||     || |||| |||| |  | |  || |||  |              ||  |
Db        605 TCCGT---GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 661

Qy        683 TTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTT 742
               |  || |     | || || | |      ||     |  | |  ||  |||||    |
Db        662 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 721

Qy        743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGA 802
               |  ||  |     |  ||||   |  |   | ||||  |||| | ||| ||  || | |
Db        722 TCCAACTTACATCTAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAA 781

Qy        803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
               |    |  ||| | | || ||  || || ||
Db        782 AATCAGATCCAAATGCGAAAGATGCTGCATTA 813


RESULT 31
ACC48531/c
ID   ACC48531 standard; DNA; 2940 BP.
XX
AC   ACC48531;
XX
DT   11-AUG-2003  (first entry)
XX
DE   Staphylococcal surface-exposed immunogenic polypeptide DNA.
XX
KW   Surface-exposed immunogenic polypeptide; SEIP; siderophore; receptor;
KW   antibacterial; vaccine; gene; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   CDS             complement(998. .1990)
FT                   /*tag=  a
FT                   /product= "SEIP"
XX
PN   WO2003020875-A2.
XX
PD   13-MAR-2003.
XX
PF   17-JUN-2002; 2002WO-US019224.
XX
PR   17-JUN-2001; 2001US-0298975P.
XX
PA   (DSQU-) D-SQUARED BIOTECHNOLOGIES INC.
XX
PI   Scott DL;
XX
DR   WPI; 2003-300870/29.
DR   P-PSDB; ABR41864.
XX
PT   Novel surface-exposed immunogenic polypeptide of Staphylococcus aureus
PT   containing receptors for siderophores or iron-binding ligands, useful for
PT   producing antibodies effective against Staphylococci infection.
XX
PS   Claim 4; Page 51-52; 62pp; English.
XX
CC   The present sequence is that of DNA encoding novel Staphylococcus aureus
CC   surface-exposed immunogenic polypeptide (SEIP) D2 SA02. To isolate SEIP
CC   genes, an expression library of S. aureus genomic DNA was screened with
CC   anti-SEIP antisera. DNA was isolated from clones expressing SEIPs, and
CC   amino acid sequences were deduced and analysed for conserved and/or
CC   functional domains. The SEIP encoded by the present DNA sequence has
CC   sequence homology with the siderophore family of periplasmic binding
CC   proteins. The invention provides methods for the identification,
CC   production and recovery of SEIPs. The SEIPs can be used individually, or
CC   in combination, to produce anti-staphylococcal antibodies useful in
CC   passive or active immunisation strategies to prevent or contain
CC   staphylococcal infection. They can also be used to develop diagnostic
CC   assays
```

```
XX
SQ   Sequence 2940 BP; 955 A; 514 C; 512 G; 959 T; 0 U; 0 Other;

  Query Match             12.3%;  Score 124;  DB 8;  Length 2940;
  Best Local Similarity   50.7%;  Pred. No. 1.6e-13;
  Matches  412;  Conservative    0;  Mismatches  380;  Indels   20;  Gaps    4;

Qy         27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
              ||||  ||   |  |||      |   | |  || |||     |||| ||||   |    ||
Db       1979 AATTAAAATGCTTGTTGTTACGCTTGCTTTCCTACTTGTTTTAGCAGGATGTAGTGGGAA 1920

Qy         87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
              |      |   | |||  | |  |   |    || ||   |  |  ||| || || | |  
Db       1919 TTCAAATAAACAATCATCTGATAACAAAGATAAGGAAACAACTTCAATTAAACATGCAAT 1860

Qy        147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
               ||||| ||  || |        | || || |||||||||   ||  |  |         
Db       1859 GGGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCAC 1800

Qy        207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
              ||| |     ||  |||||| |||||||||||| ||    |  ||       | | ||||
Db       1799 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAA 1740

Qy        267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
               |  |   |  ||    ||| | || |  |   | |    |    ||||||   |     
Db       1739 ACCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACC 1680

Qy        327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                 |||||||||||| ||||||  |||| | ||||| || |||||| | ||    ||   
Db       1679 TGCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGT 1620

Qy        387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
              |||| |  ||    ||||  |  | |||  |||||| || ||             || ||
Db       1619 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAG 1573

Qy        447 TTTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAG 505
              |||| |  || |   |  | ||| | |  ||| |   |||   |||    ||||||||||
Db       1572 TTTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAG 1513

Qy        506 GTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATA 565
              | |||||| |||||| | ||  |     |  || |  ||||| | ||| | |  | ||   
Db       1512 GGAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCC 1453

Qy        566 AAAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAAT 622
              |||||||   || | | | | ||||| ||  |  |  ||||   ||   || || ||| |
Db       1452 AAAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACT 1393

Qy        623 CAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
                 ||     || |||| |||| |  | |  || |||  |              ||  | 
Db       1392 TCCGT---GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 1336

Qy        683 TTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTT 742
               |  || |     | || || | |      ||     |  | |  ||  |||||    | 
Db       1335 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 1276

Qy        743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGA 802
               |  ||  |     |  ||||   |  |   | ||||  |||| | ||| ||  || | |
Db       1275 TCCAACTTACATCTAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAA 1216

Qy        803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
               |    |  ||| | | || ||  || || ||
Db       1215 AATCAGATCCAAATGCGAAAGATGCTGCATTA 1184


RESULT 32
AAV74549
ID   AAV74549 standard; DNA; 3775 BP.
XX
AC   AAV74549;
XX
DT   16-MAR-1999  (first entry)
XX
DE   Staphylococcus aureus contig SEQ ID #238.
```

```
XX
KW   Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW   cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW   skin infection; surgical wound infection; scalded skin syndrome;
KW   toxic shock syndrome; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1561. .1620
FT                   /*tag=  a
FT                   /note= "these bases represent a line of missing text in
FT                   the sequence listing in the specification. They are
FT                   included to maintain the nucleotide numbering given in
FT                   the specification for this DNA sequence"
FT   misc_feature    3361. .3420
FT                   /*tag=  b
FT                   /note= "these bases represent a line of missing text in
FT                   the sequence listing in the specification. They are
FT                   included to maintain the nucleotide numbering given in
FT                   the specification for this DNA sequence"
XX
PN   EP786519-A2.
XX
PD   30-JUL-1997.
XX
PF   07-JAN-1997;   97EP-00100117.
XX
PR   05-JAN-1996;   96US-0009861P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Kunsch CA,  Choi GH,  Barash SC,  Dillon PJ,  Fannon MR,  Rosen CA;
XX
DR   WPI; 1997-374922/35.
XX
PT   Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT   stored on computer readable medium and used in the production of anti-
PT   S.aureus vaccines.
XX
PS   Claim 1; Page 1021-1023; 3271pp; English.
XX
CC   This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC   of the invention. The DNA sequences are recorded on a computer readable
CC   medium, preferably selected from a floppy or hard disk, random access
CC   memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC   the S.aureus DNA sequences allows putative functions to be assigned so
CC   that protein-encoding or regulatory regions of commercial, therapeutic or
CC   industrial importance can be obtained. Specifically, sequences which are
CC   likely to encode antigens have been identified and these polypeptides can
CC   be used in a vaccine composition against S.aureus infection. The
CC   polypeptides can also be used in a kit for the immunodetection of
CC   S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC   including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC   skin and surgical wound infections, scalded skin syndrome, toxic shock
CC   syndrome, etc. Organisms transformed with the DNA sequences can be used
CC   for recombinant production of the polypeptides. The new DNA sequences
CC   (and their fragments) are useful as primers or probes for isolating
CC   homologues of any of the S.aureus DNA sequences contained on the computer
CC   readable medium
XX
SQ   Sequence 3775 BP; 1222 A; 597 C; 633 G; 1199 T; 0 U; 124 Other;

  Query Match             12.3%;  Score 124;  DB 2;  Length 3775;
  Best Local Similarity   50.7%;  Pred. No. 1.7e-13;
  Matches  412;  Conservative    0;  Mismatches  380;  Indels   20;  Gaps    4;

Qy         27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
              ||||  ||   |  |||      |   | |  || |||     |||| ||||   |    ||
Db         60 AATTAAAATGCTTGTTGTTACGCTTGCTTTCCTACTTGTTTTAGCAGGATGTAGTGGGAA 119

Qy         87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
              |      |   | |||  | |  |   |    || ||   |  |  ||| || || | |  
Db        120 TTCAAATAAACAATCATCTGATAACAAAGATAAGGAAACAACTTCAATTAAACATGCAAT 179

Qy        147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
               ||||| ||  || |        | || || |||||||||   ||  |  |
Db        180 GGGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCAC 239

Qy        207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
              ||| |     ||  |||||| |||||||||||| ||    |  ||       | | ||||
Db        240 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAA 299

Qy        267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
               |  |   |  ||    ||| | || |  |   | |    |    ||||||   |
Db        300 ACCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACC 359

Qy        327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                 |||||||||||| ||||||  |||| | |||||  || |||||| | ||    ||
Db        360 TGCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGT 419

Qy        387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
              |||| |  ||    ||||  |  | |||  |||||| || ||             || ||
Db        420 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAG 466

Qy        447 TTTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAG 505
              |||| |  || |   |  | ||| | |  ||| |   |||   |||    ||||||||||
Db        467 TTTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAG 526

Qy        506 GTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATA 565
              | |||||| |||||| | || |     |  || |  ||||| | ||| | |  | ||
Db        527 GGAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCC 586

Qy        566 AAAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAAT 622
              |||||||   || | | | | ||||| ||  |  |  ||||   ||   || || ||| |
Db        587 AAAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACT 646

Qy        623 CAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
                 ||     || |||| |||| |  | |  || |||  |              ||  |
Db        647 TCCGT---GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 703

Qy        683 TTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTT 742
               |  || |     | || || | |      ||     |  | |  ||  |||||    |
Db        704 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 763

Qy        743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGA 802
               |  ||  |     |  ||||   |  |   | ||||  |||| | ||| ||  || | |
Db        764 TCCAACTTACATCTAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAA 823

Qy        803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
               |    |  ||| | | || ||  || || ||
Db        824 AATCAGATCCAAATGCGAAAGATGCTGCATTA 855

RESULT 33
ACA22139
ID   ACA22139 standard; DNA; 796 BP.
XX
AC   ACA22139;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Prokaryotic essential gene #3796.
XX
KW   Antisense; ds; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
XX
OS   Bacillus anthracis.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,   Zamudio C,   Malone C,   Haselbeck R,   Ohlsen KL,   Zyskind JW;
PI   Wall D,   Trawick JD,   Carr GJ,   Yamamoto R,   Forsyth RA,   Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   P-PSDB; ABU18269.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 14; SEQ ID NO 10009; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential genes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 796 BP; 324 A; 124 C; 146 G; 202 T; 0 U; 0 Other;

  Query Match             11.2%;  Score 113.2;  DB 8;  Length 796;
  Best Local Similarity   57.6%;  Pred. No. 1.4e-11;
  Matches  227;  Conservative    0;  Mismatches  158;  Indels    9;  Gaps    1;

Qy        191 TTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGG 250
              |||| |    || |   || |  ||| | ||  | | |||| | || || ||||| || |
Db          2 TTGAATGGGTATATTCAGAAGACTTATTAGCACTTGGTGTTCAGCCAGTAGGGATGGCAG 61

Qy        251 ATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTT 310
              |   ||| ||  | |||       | ||  ||  || | |    | |  | |
Db         62 ACATTAAGAATTATAATAAATGGGTAAATACAAAAACAAAACCGAGTAAAGATGTTGTAG 121

Qy        311 CTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAA 370
               ||| || |||||| | ||||| |||||||||||||| ||     | |||||||||||||
Db        122 ATGTCGGGACACGTCAACAACCAAACTTAGAAGAAATTAGCCGTTTAAAACCAGATTTAA 181

Qy        371 TTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTA 430
              ||||  | | |       |    ||||  ||| | || || ||| |  ||||||| || |
Db        182 TTATCACAGCTTCATTCCGTGGTAAAGCAATTAAAAATGAATTAGAACAAATTGCACCAA 241
```

```
Qy        431 CGAT---------TGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTT 481
              |  |         |  |   | |||     | |   | | |  |||||  |  ||  | |
Db        242 CAGTTATGTTTGATCCATCAACAAGCAATAACGATCACTTTGCTGAAATGACAGAAACAT 301

Qy        482 TTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAAC 541
              |||||  |||| |||||||  | || |||||||||||||||||||||   ||||  ||
Db        302 TTAAACAAATTGCAAAAGCAGTTGGAAAAGAAGAAGAAGGTAAAAAAGTATTAGCTGATA 361

Qy        542 ACGATAAGAAAATTGAAGAATATAAAAAAGAAAT 575
                |||||   | | |  ||    |||  | ||||
Db        362 TGGATAAAGCATTCGCTGATGCAAAAGCAAAAAT 395

RESULT 34
ACF74610
ID   ACF74610 standard; DNA; 861 BP.
XX
AC   ACF74610;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Staphylococcus aureus DNA #2290.
XX
KW   Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW   enzymatic assay; antibiotic target; gene; ds.
XX
OS   Staphylococcus aureus.
XX
PN   WO200294868-A2.
XX
PD   28-NOV-2002.
XX
PF   27-MAR-2002; 2002WO-IB002637.
XX
PR   27-MAR-2001; 2001GB-00007661.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Masignani V,  Mora M,  Scarselli M;
XX
DR   WPI; 2003-120786/11.
DR   P-PSDB; ABM73050.
XX
PT   New Staphylococcus aureus protein, useful as a vaccine for treating or
PT   preventing Staphylococcal infection, specifically an infection caused by
PT   S. aureus, e.g. sepsis.
XX
PS   Claim 6; SEQ ID NO 4579; 49pp; English.
XX
CC   The invention relates to novel genes and encoded proteins from
CC   Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC   nucleic acid encoding the protein, or an antibody to the protein, is
CC   useful as a pharmaceutical, particularly as a vaccine for treating or
CC   preventing infection due to Staphylococcus bacteria, specifically an
CC   infection caused by S. aureus. The composition is particularly useful for
CC   treating or preventing sepsis in a patient. The composition can also be
CC   used for diagnostics. The protein is also used in an assay for enzymatic
CC   studies and as a target for antibiotics. This sequence represents one of
CC   the novel S. aureus genes of the invention
XX
SQ   Sequence 861 BP; 348 A; 123 C; 158 G; 232 T; 0 U; 0 Other;

  Query Match             10.9%;  Score 110.2;  DB 8;  Length 861;
  Best Local Similarity   51.8%;  Pred. No. 5.1e-11;
  Matches  358;  Conservative    0;  Mismatches  313;  Indels   20;  Gaps    4;

Qy        148 GGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGTT 207
              ||||| ||  || |        | || || |||||||||   ||  |  |         |
Db          4 GGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCACT 63

Qy        208 GATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAAT 267
              || |     ||  |||||| |||||||||||| ||    |  ||       | | ||||
Db         64 GACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAAA 123
```

```
Qy        268 CGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAG 327
              |  |   |  ||    ||| | || |  |   | |    |    ||||||   |
Db        124 CCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACCT 183

Qy        328 CAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAAT 387
                |||||||||||| ||||||  |||| | ||||| || |||||| | ||    ||   |
Db        184 GCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGTT 243

Qy        388 AGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGT 447
              ||| |  ||    ||||  |  | |||  |||||| || ||             || |||
Db        244 AGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAGT 290

Qy        448 TTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGG 506
              ||| |  || |   |  | ||| | |  ||| |   |||   |||    |||||||||||
Db        291 TTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAGG 350

Qy        507 TAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAA 566
               |||||| |||||| | || |     |  || |  ||||| | ||| | |  | ||   |
Db        351 GAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCCA 410

Qy        567 AAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATC 623
              ||||||   || | | | | ||||| ||  |  |  ||||   ||   || || ||| |
Db        411 AAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACTT 470

Qy        624 AGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTT 683
                ||     || |||| |||| |  | |  || |||  |              ||  |
Db        471 CCGTGC---TGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATGA 527

Qy        684 TAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTA 743
              |  || |     | || || | |      ||     |  | |  ||  |||||    |
Db        528 TTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTAT 587

Qy        744 CTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGAC 803
              |  ||  |     |  ||||   |  |   | ||||  |||| | ||| ||  || | |
Db        588 CCAACTTACATCTAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAAA 647

Qy        804 AAACAAAGCAAGTTCTAACGAACCTTCACTA 834
              |    |  ||| | | || ||  || || ||
Db        648 ATCAGATCCAAATGCGAAAGATGCTGCATTA 678

RESULT 35
AAH54418
ID   AAH54418 standard; DNA; 3046 BP.
XX
AC   AAH54418;
XX
DT   03-SEP-2001  (first entry)
XX
DE   S. epidermidis genomic polynucleotide sequence SEQ ID NO:3782.
XX
KW   Staphylococcus epidermidis SR1 strain; infection; diagnosis; vaccination;
KW   endocarditis; ds.
XX
OS   Staphylococcus epidermidis.
XX
PN   WO200134809-A2.
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US030782.
XX
PR   09-NOV-1999;   99US-0164258P.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Kimmerly WJ;
XX
DR   WPI; 2001-316495/33.
XX
PT   Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT   useful for vaccinating against infections, e.g. endocarditis.
```

```
XX
PS   Claim 8; Page 1413-1414; 2188pp; English.
XX
CC   AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC   (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)
CC   and (II) can have antibacterial activity and therefore can be used in
CC   vaccination. The nucleic acids (I) may be used to produce the S.
CC   epidermidis polypeptides (II) via the production of vectors containing
CC   them which are used to produce hosts cells which express the
CC   polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC   used to vaccinate subjects and to raise antibodies against the bacteria.
CC   The polypeptides may also be used to assay for other inhibitors of their
CC   activity and therefore identify compounds that may be used for the
CC   treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC   AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC   polynucleotide sequences from the present invention. AAH55091 to AAH55098
CC   represent oligonucleotide sequences and primers which are used in the
CC   exemplification of the present invention. N.B. The present invention
CC   specifically claims all the polynucleotide sequences given in the
CC   sequence listing of the present specification, however the sequence
CC   listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC   in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
CC   for SEQ ID NO:4455 to 4464
XX
SQ   Sequence 3046 BP; 1103 A; 468 C; 436 G; 1039 T; 0 U; 0 Other;

  Query Match             10.9%;  Score 109.4;  DB 4;  Length 3046;
  Best Local Similarity   95.0%;  Pred. No. 7.7e-11;
  Matches  113;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 GGAGTGGAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2705 GGAGTGGAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTA 2764

Qy         61 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAA 119
              ||||||||||||||||||||||||||||||||||||||||||||||| |    || |||
Db       2765 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAATGTAATATAAAA 2823


RESULT 36
ACA32000
ID   ACA32000 standard; DNA; 900 BP.
XX
AC   ACA32000;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Prokaryotic essential gene #13657.
XX
KW   Antisense; ds; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
XX
OS   Enterobacter cloacae.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,   Zamudio C,   Malone C,   Haselbeck R,   Ohlsen KL,   Zyskind JW;
PI   Wall D,   Trawick JD,   Carr GJ,   Yamamoto R,   Forsyth RA,   Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   P-PSDB; ABU28130.
```

```
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 14; SEQ ID NO 19870; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential genes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 900 BP; 198 A; 266 C; 256 G; 180 T; 0 U; 0 Other;

  Query Match             10.2%;  Score 103.2;  DB 8;  Length 900;
  Best Local Similarity   52.9%;  Pred. No. 9.7e-10;
  Matches  222;  Conservative    0;  Mismatches  198;  Indels    0;  Gaps    0;

Qy        137 AGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGT 196
              || ||||| | || ||    | | |   ||||   ||  ||||  |||| || || ||
Db         74 AGGACGAACACGGCACGTTTACACTCGATAAAACGCCACAACGGATTGTGGTGCTGGAAC 133

Qy        197 ATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
                || || |  |||||| | |  ||  | || || |  ||  | || || || || ||||
Db        134 TCTCGTTCGCCGATGCGCTGGCCGCCGTGGACGTCAGCCCGATCGGTATTGCCGACGATA 193

Qy        257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
              || |   ||| || ||  |      | |  | |   |  |   |   |     || || |
Db        194 ACGATGCAAAACGCATCCTGCCCGAAGTGCGTGCGCACCTGAAACCGTGGCAGTCCGTCG 253

Qy        317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
              |||| ||   ||| || | | | ||||  ||   |   ||  |||||||| | || ||||
Db        254 GAACGCGCGCGCAGCCGAGCCTGGAAGCCATTGCTGCCCTGAAACCAGACCTGATCATTG 313

Qy        377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
              | || |  | | | ||    || ||||| |  | |||  |  |||| || ||     |
Db        314 CCGACAGCAGTCGCCATGCGGGGATTTACACTGCCTTGCAGCAAATCGCGCCGGTACTGC 373

Qy        437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
                || ||      | | | |   ||   ||||||| |  |  |       |  ||
Db        374 TGCTTAAGTCCCGCAACGAAACCTACGCTGAAAATTTGCACTCAGCGGCTATCATCGGCG 433

Qy        497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
              |||   ||||||||| ||  |||     |    ||  | ||| ||||  |  |||| || |
Db        434 AAGTGGTAGGTAAAAAACGAGAGATGCAGGCGCGTCTGGAACAACATAAAGAGAAGATGG 493
```

```
RESULT 37
ACF70890
ID   ACF70890 standard; DNA; 909 BP.
XX
AC   ACF70890;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Photorhabdus luminescens nucleotide sequence #9357.
XX
KW   Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
KW   detection; food; gene expression; plant; animal; microorganism; toxin;
KW   antibiotic; biopesticide; virulence factor; disease model; plague;
KW   whooping cough; gene; ds.
XX
OS   Photorhabdus luminescens.
XX
PN   WO200294867-A2.
XX
PD   28-NOV-2002.
XX
PF   07-FEB-2002; 2002WO-IB003040.
XX
PR   07-FEB-2001; 2001FR-00001659.
XX
PA   (INSP ) INST PASTEUR.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Duchaud E,  Taourit S,  Glaser P,  Frangeul L,  Kunst F,  Danchin A;
PI   Buchrieser C;
XX
DR   WPI; 2003-148459/14.
XX
PT   Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
PT   useful e.g. as therapeutic antimicrobials and agricultural pesticides.
XX
PS   Claim 2; SEQ ID NO 9357; 1205pp; French.
XX
CC   The invention relates to the isolation of genes and their encoded
CC   proteins from Photorhabdus luminescens. The isolated sequences are
CC   sources of probes and primers for detecting the genome of P. luminescens
CC   and related species; to study polymorphisms; for gene analysis and for
CC   detection/amplification of the genes. Antibodies (Ab) raised against the
CC   polypeptides encoded by the genes are used for detection/identification
CC   of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
CC   carry a gene-containing vector are used to select compounds that
CC   modulate, regulate, induce or inhibit expression of the genes in plants,
CC   animals or microorganisms other than P. luminescens and are able to alter
CC   response or sensitivity to toxins and antibiotics produced by P.
CC   luminescens. Cells transformed to express the genes are useful for
CC   recombinant production of the proteins, particularly toxins and
CC   antibacterials useful as insecticides, bactericides and fungicides. The
CC   genes, proteins, vectors containing the genes and Ab are also useful
CC   therapeutically (to treat microbial infection by bacteria or fungi that
CC   are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC   biopesticides. Other uses of the genes and the proteins are as virulence
CC   factors and for identifying targets of human diseases for which P.
CC   luminescens is a model (particularly plague and whooping cough). This
CC   sequence represents one of the isolated P. luminescens genes
XX
SQ   Sequence 909 BP; 230 A; 258 C; 225 G; 196 T; 0 U; 0 Other;

  Query Match             9.6%;  Score 96.8;  DB 10;  Length 909;
  Best Local Similarity   51.6%;  Pred. No. 1.4e-08;
  Matches  221;  Conservative    0;  Mismatches  207;  Indels    0;  Gaps    0;

Qy        137 AGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGT 196
              || | ||  | || | |   |   |   ||     ||  | || || || || || || |
Db         80 AGGATGAGCAGGGCAGTTTTACCCTCAATACCGTTCCCCAGCGCGTAGTCGTGCTGGAAT 139

Qy        197 ATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
                |||||||  ||||| || |  ||  |  || | |  || || || || || |||||||
Db        140 TGTCATTTGCCGATGCATTGGCCGCCATCAATATCAGTCCGGTAGGTATTGCCGATGATA 199

Qy        257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
              || |     | || ||| | |      |  |  |    |||      |     ||    |
Db        200 ACGATCCGCAGCGCATTCTCACCGATGTTCGCCAGCGTATTAAGCCCTGGCAATCAACCG 259

Qy        317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
              | || ||    ||||| | | | ||||  |||||    || ||||| ||| | || ||||
Db        260 GCACTCGCGCCCAACCAAGCCTTGAAGCCATCAGCGCGCTGAAACCTGATCTCATCATTG 319

Qy        377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
              | || | | |  | |||   || ||||||   |   | ||    || ||||||    |
Db        320 CAGACAGTCAACGCCACGCGGGGATTTATCGGGCACTCAAGGGCATCGCTCCTGTCCTGC 379

Qy        437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
                || |||     | | | |   ||  | || ||  |  |  ||      |  ||    |
Db        380 TGCTAAAATCCCGCAACGAAACCTACGAAGAGAACCTGCAATCTGCCGCGATCATCGGCA 439

Qy        497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
              |||   | || |||||    ||     | |||||  | | ||| |||  | |    ||
Db        440 AAGTCATGGGGAAAGAGAGTGAGATGCAGAAACGGCTGGCAGAGCACCGTGAACGCATGA 499

Qy        557 AAGAATAT 564
              |||  |||
Db        500 AAGGCTAT 507


RESULT 38
ADR20887/c
ID   ADR20887 standard; DNA; 25860 BP.
XX
AC   ADR20887;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Photorhabdus luminescens genomic DNA encoding two insecticide proteins.
XX
KW   insecticide protein; transgenic plant; insect-resistant plant; wheat;
KW   maize; gene; ds.
XX
OS   Photorhabdus luminescens subsp. laumondii.
XX
FH   Key             Location/Qualifiers
FT   CDS             20872. .21306
FT                   /*tag=  a
FT                   /product= "Photorhabdus luminescens insecticide protein
FT                   #1"
FT   CDS             21345. .22598
FT                   /*tag=  b
FT                   /product= "Photorhabdus luminescens insecticide protein
FT                   #2"
XX
PN   WO2003087377-A1.
XX
PD   23-OCT-2003.
XX
PF   17-APR-2003; 2003WO-FR001239.
XX
PR   17-APR-2002; 2002FR-00004798.
XX
PA   (INSP ) INST PASTEUR.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX
DR   WPI; 2003-833739/77.
DR   P-PSDB; ADR20889, ADR20891.
XX
PT   New nucleic acid from Photorhabdus luminescens, useful for producing
PT   insecticidal polypeptides active against Lepidoptera, and for producing
PT   insect resistant transgenic plants.
XX
PS   Claim 22; SEQ ID NO 1; 44pp; French.
XX
```

```
CC   The invention comprises the amino acid and coding sequences of proteins
CC   from Photorhabdus luminescens which are active against insects. The
CC   proteins of the invention are toxic to Lepidoptera, specifically the
CC   genera Plutella, Heliothis, Helicoverpa, Spodoptera and Ostrinia. The DNA
CC   and protein sequences of the invention are useful in the preparation of
CC   transgenic, insect-resistant plants, specifically wheat and maize. The
CC   present Photorhabdus luminescens genomic DNA contains the coding sequence
CC   for the two insecticide proteins of the invention.
XX
SQ   Sequence 25860 BP; 7315 A; 5392 C; 6047 G; 7106 T; 0 U; 0 Other;

  Query Match             9.6%;  Score 96.8;  DB 11;  Length 25860;
  Best Local Similarity   51.6%;  Pred. No. 1.8e-08;
  Matches  221;  Conservative    0;  Mismatches  207;  Indels    0;  Gaps    0;

Qy        137 AGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGT 196
              || | ||  | || | |   |   |   ||     ||  | || || || || || || |
Db       8497 AGGATGAGCAGGGCAGTTTTACCCTCAATACCGTTCCCCAGCGCGTAGTCGTGCTGGAAT 8438

Qy        197 ATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
                |||||||  ||||| || |  ||  |  || | |  || || || || || |||||||
Db       8437 TGTCATTTGCCGATGCATTGGCCGCCATCAATATCAGTCCGGTAGGTATTGCCGATGATA 8378

Qy        257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
              || |     | || ||| | |      |  |  |    |||      |     ||    |
Db       8377 ACGATCCGCAGCGCATTCTCACCGATGTTCGCCAGCGTATTAAGCCCTGGCAATCAACCG 8318

Qy        317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
              | || ||     ||||| | | | ||||  |||||    || ||||| ||| | || ||||
Db       8317 GCACTCGCGCCCAACCAAGCCTTGAAGCCATCAGCGCGCTGAAACCTGATCTCATCATTG 8258

Qy        377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
              | || | | |  | |||   || ||||||   |   | ||    || ||||||    |  
Db       8257 CAGACAGTCAACGCCACGCGGGGATTTATCGGGCACTCAAGGGCATCGCTCCTGTCCTGC 8198

Qy        437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
                || |||     | | | |   ||  | || ||  |  |  ||      |  ||    |
Db       8197 TGCTAAAATCCCGCAACGAAACCTACGAAGAGAACCTGCAATCTGCCGCGATCATCGGCA 8138

Qy        497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
              |||   | || |||||    ||     | |||||  | | ||| |||  | |    ||  
Db       8137 AAGTCATGGGGAAAGAGAGTGAGATGCAGAAACGGCTGGCAGAGCACCGTGAACGCATGA 8078

Qy        557 AAGAATAT 564
              |||  |||
Db       8077 AAGGCTAT 8070


RESULT 39
ACF67367_39/c
Continuation (40 of 57) of ACF67367 from base 3900001 (Photorhabdus luminescens nucleoti
WP Sequence split into 57 fragments  LOCUS ACF67367 Accession Acf67367
WP     Fragment Name      Begin      End
WP     ACF67367_00            1      110000
WP     ACF67367_01       100001      210000
WP     ACF67367_02       200001      310000
WP     ACF67367_03       300001      410000
WP     ACF67367_04       400001      510000
WP     ACF67367_05       500001      610000
WP     ACF67367_06       600001      710000
WP     ACF67367_07       700001      810000
WP     ACF67367_08       800001      910000
WP     ACF67367_09       900001     1010000
WP     ACF67367_10      1000001     1110000
WP     ACF67367_11      1100001     1210000
WP     ACF67367_12      1200001     1310000
WP     ACF67367_13      1300001     1410000
WP     ACF67367_14      1400001     1510000
WP     ACF67367_15      1500001     1610000
WP     ACF67367_16      1600001     1710000
WP     ACF67367_17      1700001     1810000
WP     ACF67367_18      1800001     1910000
WP     ACF67367_19      1900001     2010000
WP     ACF67367_20      2000001     2110000
WP     ACF67367_21      2100001     2210000
WP     ACF67367_22      2200001     2310000
WP     ACF67367_23      2300001     2410000
WP     ACF67367_24      2400001     2510000
WP     ACF67367_25      2500001     2610000
WP     ACF67367_26      2600001     2710000
WP     ACF67367_27      2700001     2810000
WP     ACF67367_28      2800001     2910000
WP     ACF67367_29      2900001     3010000
WP     ACF67367_30      3000001     3110000
WP     ACF67367_31      3100001     3210000
WP     ACF67367_32      3200001     3310000
WP     ACF67367_33      3300001     3410000
WP     ACF67367_34      3400001     3510000
WP     ACF67367_35      3500001     3610000
WP     ACF67367_36      3600001     3710000
WP     ACF67367_37      3700001     3810000
WP     ACF67367_38      3800001     3910000
WP     ACF67367_39      3900001     4010000
WP     ACF67367_40      4000001     4110000
WP     ACF67367_41      4100001     4210000
WP     ACF67367_42      4200001     4310000
WP     ACF67367_43      4300001     4410000
WP     ACF67367_44      4400001     4510000
WP     ACF67367_45      4500001     4610000
WP     ACF67367_46      4600001     4710000
WP     ACF67367_47      4700001     4810000
WP     ACF67367_48      4800001     4910000
WP     ACF67367_49      4900001     5010000
WP     ACF67367_50      5000001     5110000
WP     ACF67367_51      5100001     5210000
WP     ACF67367_52      5200001     5310000
WP     ACF67367_53      5300001     5410000
WP     ACF67367_54      5400001     5510000
WP     ACF67367_55      5500001     5610000
WP     ACF67367_56      5600001     5648894

  Query Match             9.6%;  Score 96.8;  DB 10;  Length 110000;
  Best Local Similarity   51.6%;  Pred. No. 1.9e-08;
  Matches  221;  Conservative    0;  Mismatches  207;  Indels    0;  Gaps    0;

Qy        137 AGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGT 196
              || | ||  | || | |   |   |   ||     ||  | || || || || || || |
Db      95107 AGGATGAGCAGGGCAGTTTTACCCTCAATACCGTTCCCCAGCGCGTAGTCGTGCTGGAAT 95048

Qy        197 ATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
                |||||||  ||||| || |  ||  |  || | |  || || || || || |||||||
Db      95047 TGTCATTTGCCGATGCATTGGCCGCCATCAATATCAGTCCGGTAGGTATTGCCGATGATA 94988

Qy        257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
              || |     | || ||| | |      |  |  |    |||      |     ||    |
Db      94987 ACGATCCGCAGCGCATTCTCACCGATGTTCGCCAGCGTATTAAGCCCTGGCAATCAACCG 94928

Qy        317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
              | || ||     ||||| | | | ||||  |||||    || ||||| ||| | || ||||
Db      94927 GCACTCGCGCCCAACCAAGCCTTGAAGCCATCAGCGCGCTGAAACCTGATCTCATCATTG 94868

Qy        377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
              | || | | |  | |||   || ||||||   |   | ||    || ||||||    |  
Db      94867 CAGACAGTCAACGCCACGCGGGGATTTATCGGGCACTCAAGGGCATCGCTCCTGTCCTGC 94808

Qy        437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
                || |||     | | | |   ||  | || ||  |  |  ||      |  ||    |
Db      94807 TGCTAAAATCCCGCAACGAAACCTACGAAGAGAACCTGCAATCTGCCGCGATCATCGGCA 94748

Qy        497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
              |||   | || |||||    ||     | |||||  | | ||| |||  | |    ||  
Db      94747 AAGTCATGGGGAAAGAGAGTGAGATGCAGAAACGGCTGGCAGAGCACCGTGAACGCATGA 94688

Qy        557 AAGAATAT 564
              |||  |||
Db      94687 AAGGCTAT 94680
```

```
RESULT 40
ACF65388_08
Continuation (9 of 13) of ACF65388 from base 800001 (Photorhabdus luminescens nucleotide
WP Sequence split into 13 fragments  LOCUS ACF65388 Accession Acf65388
WP     Fragment Name       Begin      End
WP     ACF65388_00             1    110000
WP     ACF65388_01        100001    210000
WP     ACF65388_02        200001    310000
WP     ACF65388_03        300001    410000
WP     ACF65388_04        400001    510000
WP     ACF65388_05        500001    610000
WP     ACF65388_06        600001    710000
WP     ACF65388_07        700001    810000
WP     ACF65388_08        800001    910000
WP     ACF65388_09        900001   1010000
WP     ACF65388_10       1000001   1110000
WP     ACF65388_11       1100001   1210000
WP     ACF65388_12       1200001   1225559

  Query Match             9.6%;  Score 96.8;  DB 10;  Length 110000;
  Best Local Similarity   51.6%;  Pred. No. 1.9e-08;
  Matches  221;  Conservative    0;  Mismatches  207;  Indels    0;  Gaps    0;

Qy        137 AGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGT 196
              || | ||  | || | |   |   |   ||     ||  | || || || || || || |
Db      29830 AGGATGAGCAGGGCAGTTTTACCCTCAATACCGTTCCCCAGCGCGTAGTCGTGCTGGAAT 29889

Qy        197 ATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
                |||||||  ||||| || |  ||  |  || | |  || || || || || |||||||
Db      29890 TGTCATTTGCCGATGCATTGGCCGCCATCAATATCAGTCCGGTAGGTATTGCCGATGATA 29949

Qy        257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
              || |     | || ||| | |      |  |  |    |||      |     ||    |
Db      29950 ACGATCCGCAGCGCATTCTCACCGATGTTCGCCAGCGTATTAAGCCCTGGCAATCAACCG 30009

Qy        317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
              | || ||    ||||| | | | ||||  |||||    || ||||| ||| | || ||||
Db      30010 GCACTCGCGCCCAACCAAGCCTTGAAGCCATCAGCGCGCTGAAACCTGATCTCATCATTG 30069

Qy        377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
              | || | | |  | |||   || ||||||   |   | ||    || ||||||    |
Db      30070 CAGACAGTCAACGCCACGCGGGGATTTATCGGGCACTCAAGGGCATCGCTCCTGTCCTGC 30129

Qy        437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
                || |||     | | | |   ||  | || ||  |  |  ||      |  ||    |
Db      30130 TGCTAAAATCCCGCAACGAAACCTACGAAGAGAACCTGCAATCTGCCGCGATCATCGGCA 30189

Qy        497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
              |||   | || |||||    ||     | ||||||  | | ||| |||  | |    ||
Db      30190 AAGTCATGGGGAAAGAGAGTGAGATGCAGAAACGGCTGGCAGAGCACCGTGAACGCATGA 30249

Qy        557 AAGAATAT 564
              |||  |||
Db      30250 AAGGCTAT 30257

Search completed: August 26, 2005, 04:36:53
Job time : 653 secs
```

4

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2005  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         August 26, 2005, 03:38:21 ; Search time 3906 Seconds
                                            (without alignments)
                                            9823.037 Million cell updates/sec

Title:          US-10-724-972A-2580
Perfect score:  1008
Sequence:       1 ggagtggaatcagtgagagg..........gtaaaaaagataataagtaa 1008

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters:      68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       EST:*
                1:  gb_est1:*
                2:  gb_est2:*
                3:  gb_htc:*
                4:  gb_est3:*
                5:  gb_est4:*
                6:  gb_est5:*
                7:  gb_est6:*
                8:  gb_gss1:*
                9:  gb_gss2:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                         %
Result                 Query
   No.      Score      Match Length DB  ID                       Description
-----------------------------------------------------------------------------------
       1     96.8       9.6     801  9  CL674627                 CL674627 PRI0112c_
 c     2     95.2       9.4     625  7  CF307287                 CF307287 HDA1--06-
 c     3     87.4       8.7    1566  9  CG757757                 CG757757 P053-1-D0
       4     87.4       8.7    1632  9  CL082569                 CL082569 CH216-167
       5     86.6       8.6    1135  9  CNS033GQ                 AL226115 Tetraodon
       6     86.4       8.6    1101  9  CNS0021J                 AL061936 Drosophil
 c     7     83.6       8.3     773  8  BH651171                 BH651171 BOMJP20TR
 c     8     83.6       8.3    1172  3  CR734084                 CR734084 Tetraodon
 c     9     82.4       8.2    1362  9  AG429915                 AG429915 Mus muscu
      10     81.6       8.1    1042  9  CL461494                 CL461494 SAIL_1148
      11     81.4       8.1    1017  9  CL101048                 CL101048 ISB1-3604
      12     81.4       8.1    1626  7  CF238805                 CF238805 AGENCOURT
      13     81.2       8.1    1594  9  CL038406                 CL038406 CH216-46A
 c    14     80.8       8.0    1101  9  CNS00HX9                 AL073856 Drosophil
 c    15     80.8       8.0    1519  9  AG386893                 AG386893 Mus muscu
      16     80.6       8.0    1402  9  CL078552                 CL078552 CH216-151
 c    17     80.4       8.0    1506  9  AG278469                 AG278469 Mus muscu
 c    18     80.2       8.0    1225  9  CG745927                 CG745927 P039-1-A0
      19       80       7.9    1256  9  CL119201                 CL119201 ISB1-76J1
 c    20     79.6       7.9     776  9  CL694436                 CL694436 PRI0164b_
 c    21     79.6       7.9    1231  9  CG755650                 CG755650 P051-1-G0
 c    22     79.6       7.9    1251  9  AG332167                 AG332167 Mus muscu
      23     79.2       7.9     625  9  CNS036A2                 AL229763 Tetraodon
 c    24     79.2       7.9    1101  9  CNS017Y0                 AL108642 Drosophil
```

```
 c    25     79.2       7.9    1353  9  CG744812                 CG744812 P037-3-B0
 c    26       79       7.8    1392  9  CG757503                 CG757503 P052-4-C0
      27       79       7.8    1536  9  CL078538                 CL078538 CH216-151
      28     78.8       7.8    1162  9  CL077122                 CL077122 CH216-143
 c    29     78.8       7.8    1185  2  BF273407                 BF273407 GA__Eb001
 c    30     78.8       7.8    1217  9  CL078190                 CL078190 CH216-148
 c    31     78.8       7.8    1276  9  CG754010                 CG754010 P049-2-C0
 c    32     78.8       7.8    1472  9  AG350118                 AG350118 Mus muscu
 c    33     78.6       7.8    1352  8  CC268594                 CC268594 CH261-67F
 c    34     78.6       7.8    1811  9  CG753732                 CG753732 P048-4-G0
      35     78.4       7.8    1163  9  CL081338                 CL081338 CH216-161
 c    36     78.4       7.8    1297  9  CG758143                 CG758143 P053-3-B1
 c    37     78.4       7.8    1380  9  CG744815                 CG744815 P037-3-B0
 c    38     78.4       7.8    1594  9  CL110653                 CL110653 ISB1-53P2
      39     78.2       7.8     971  6  CD388253                 CD388253 AGENCOURT
 c    40     78.2       7.8    1110  9  CL037168                 CL037168 CH216-42N
 c    41     78.2       7.8    1277  9  AG288477                 AG288477 Mus muscu
 c    42     78.2       7.8    1373  9  CG750869                 CG750869 P045-2-E1
 c    43     78.2       7.8    1453  9  AJ591978                 AJ591978 Arabidops
 c    44       78       7.7     722  7  CV481731                 CV481731 AGENCOURT
 c    45       78       7.7     964  9  CNS07E8R                 AL441457 T7 end of


                                  ALIGNMENTS


RESULT 1
CL674627
LOCUS       CL674627                801 bp    DNA     linear   GSS 09-JUL-2004
DEFINITION  PRI0112c_G03_2 - PRI0112c.BR (801) Mixed stage fosmid library of P.
            pacificus var. California Pristionchus pacificus genomic, genomic
            survey sequence.
ACCESSION   CL674627
VERSION     CL674627.1  GI:50177972
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 801)
  AUTHORS   Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
  TITLE     AppaDB: an AcedB database for the nematode satellite organism
            Pristionchus pacificus
  JOURNAL   Nucleic Acids Res. 32 (1), D421-D422 (2004)
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            This library was generated at Caltech, Pasadena, USA and end
            sequenced at Vancouver, Canada.
            Seq primer: T7
            Class: fosmid ends.
FEATURES             Location/Qualifiers
     source          1..801
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Mixed stage fosmid library of P. pacificus
                     var. California"
                     /note="Vector: pEpifos-5 Fosmid vector"
ORIGIN

  Query Match             9.6%;  Score 96.8;  DB 9;  Length 801;
  Best Local Similarity   51.9%;  Pred. No. 3.3e-09;
  Matches  218;  Conservative    0;  Mismatches  202;  Indels    0;  Gaps    0;

Qy        137 AGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGT 196
              || ||||| | || ||    | | |    |||   ||  ||||  |||| || || ||  
Db         83 AGGACGAACACGGCACGTTTACACTCGAAAAAACGCCACAACGGATTGTGGTGCTGGAAC 142
```

```
Qy        197 ATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
                || || |  |||||| | |  ||  | || || |  ||  | || || || || ||||
Db        143 TCTCGTTCGCCGATGCGCTGGCCGCCGTGGACGTCAGCCCGATCGGTATTGCCGACGATA 202

Qy        257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
              || |   ||| || ||  |      | |  | |   |  |   |   |     || || |
Db        203 ACGATGCAAAACGCATCCTGCCCGAAGTGCGTGCGCACCTGAAACCGTGGCAGTCCGTCG 262

Qy        317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
              |||| ||   ||| || | | | ||||  ||       || ||||||||  | || ||||
Db        263 GAACGCGCGCGCAGCCGAGCCTGGAAGCCATTGCCGCTCTGAAACCAGACCTGATCATTG 322

Qy        377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
              | || |  | | | ||    ||  |||| |  | |||  |  |||| || ||     |
Db        323 CCGACAGCAGTCGCCATGCGGGGGTTTACATCGCCTTGCAGCAAATCGCGCCGGTACTGC 382

Qy        437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
                || ||      | | | |   ||   ||||||| |  |  ||      |  ||
Db        383 TGCTTAAGTCCCGCAACGAAACCTACGCTGAAAATTTGCAATCTGCGGCTATCATCGGCG 442

Qy        497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
              ||    | |||||| |   |||     |   |||| | ||| ||||  |  |||  || |
Db        443 AAATGGTGGGTAAAAAGCGAGAGATGCAGGCACGCCTGGAACAACATAAAGAGAGGATGG 502


RESULT 2
CF307287/c
LOCUS       CF307287                 625 bp    mRNA    linear   EST 15-AUG-2003
DEFINITION  HDA1--06-E11.g1 OsHDAC1-overexpressing transgenic rice lambda phage
            cDNA library I (HDA1) Oryza sativa (japonica cultivar-group) cDNA
            clone HDA1--06-E11, mRNA sequence.
ACCESSION   CF307287
VERSION     CF307287.1  GI:33679048
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 625)
  AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
            Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
  TITLE     Large-scale Sequencing Analysis of Rice ESTs
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Nahm B.H.
            Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
            of Bioscience and Bioinformatics, MyongJi University
            Yongin, Kyeonggi, Korea
            Tel: 82 31 330 6193
            Fax: 82 31 321 6355
            Email: bhnahm@ggbio.com, bhnahm@bio.myongji.ac.kr.
FEATURES             Location/Qualifiers
     source          1..625
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nackdong"
                     /db_xref="taxon:39947"
                     /clone="HDA1--06-E11"
                     /tissue_type="callus"
                     /dev_stage="proliferated callus on 2N6 media for 2 weeks"
                     /lab_host="E.coli SOLR"
                     /clone_lib="OsHDAC1-overexpressing transgenic rice lambda
                     phage cDNA library I (HDA1)"
                     /note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
                     XhoI; Callus was treated with ABA(20um) for 1hour. cDNA
                     was inserted into lambda Uni-ZAP XR vector at 5' end with
                     EcoRI and 3' end with XhoI site. mRNA was derived from
                     rice Histone Deacetylase overexpression line."
ORIGIN

  Query Match             9.4%;  Score 95.2;  DB 7;  Length 625;
  Best Local Similarity   51.7%;  Pred. No. 6.9e-09;
```

```
  Matches  217;  Conservative    0;  Mismatches  203;  Indels    0;  Gaps    0;

Qy        137 AGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGT 196
              || ||||| | || ||    | | |    |||   ||  ||||  |||| || || ||
Db        615 AGGACGAACACGGCACGTTTACACTCGAAAAAACGCCACAACGGATTGTGGTGCTGGAAC 556

Qy        197 ATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
                || || |  |||||| | |  ||  | || || |  ||  | || || || || ||||
Db        555 TCTCGTTCGCCGATGCGCTGGCCGCCGTGGACGTCATCCCGATCGGTATTGCCGACGATA 496

Qy        257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
              || |   ||| || ||  |      | |  | |   |  |   |   |     || || |
Db        495 ACGATGCAAAACGCATCCTGCCCGAAGTGCGTGCGCACCTGAAACCGTGGCAGTCCGTCG 436

Qy        317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
              |||| ||   ||| || | | | ||||  ||       || ||||||||  | || ||||
Db        435 GAACGCGCGCGCAGCCGAGCCTGGAAGCCATTGCCGCTCTGAAACCAGACCTGATCATTG 376

Qy        377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
              | || |  | | | ||    ||  |||| |  | |||  |  |||| || ||     |
Db        375 CCGACAGCAGTCGCCATGCGGGGGTTTACATCGCCTTGCAGCAAATCGCGCCGGTACTGC 316

Qy        437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
                || ||      | | | |   ||   ||||||| |  |  ||      |  ||
Db        315 TGCTTAAGTCCCGCAACGAAACCTACGCTGAAAATTTGCAATCTGCGGCTATCATCGGCG 256

Qy        497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
              ||    | |||||| |   |||     |   |||  | ||| ||||  |  |||  || |
Db        255 AAATGGTGGGTAAAAAGCGAGAGATGCAGGCACGTCTGGAACAACATAAAGAGAGGATGG 196


RESULT 3
CG757757/c
LOCUS       CG757757                1566 bp    DNA     linear   GSS 24-OCT-2003
DEFINITION  P053-1-D07.zc Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG757757
VERSION     CG757757.1  GI:37986636
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 1566)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
  TITLE     An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
  MEDLINE   22835951
   PUBMED   12884007
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..1566
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest of
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector."
ORIGIN

  Query Match             8.7%;  Score 87.4;  DB 9;  Length 1566;
  Best Local Similarity   38.3%;  Pred. No. 2.6e-07;
  Matches  289;  Conservative    0;  Mismatches  466;  Indels    0;  Gaps    0;

Qy        254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
              | || ||||||||    |  |  |||  |  || | |  |||    |||| | |
Db       1528 AAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAANAAAANAAAAAAAAAAAAAAAANANAN 1469

Qy        314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
               |     |   ||  ||   ||     || |||  |  |||   |||  | |   |
Db       1468 AANANNAAANAAAAAAAANANAAANNAAAAAAAAAAAAAAAAAAAAANAAAAAAAANAAN 1409

Qy        374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                    | || || | | | ||    |   | | | |   ||| ||||        |  |
Db       1408 AAANNAAAAANAAAANAAAAAANANAAAAAAAANANAAAAAAAAAAAAAAANNAAAAAAA 1349

Qy        434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                 ||   ||||      |   |    |  |  |||| |   |       |||| ||
Db       1348 AAAAANAAAAAAAAANNAAAAAAANAAAANAAAAAAAAANNAAAAAAAANAAAAAAAANA 1289

Qy        494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
               ||||     |   ||| || || ||    |||||       || ||    | || ||||
Db       1288 AAAAAANANNANAAAAAAAAAAAAAAAANNAAAAAAANNNNAAAAAANNAAAAAAAAAAA 1229

Qy        554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
                 || |    |||||| ||| |   |   | |||||  | ||   |        |  |
Db       1228 AAAAANANANAAAAAAAAAAAAANAAAANAAAAAAAAAAANAAAAAAAANANANNAANAA 1169

Qy        614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
                    ||  |             |   ||   |              ||     ||
Db       1168 AAAAANAAAAAAAAAAAANNNANAAANAAAAANANNAAANNNANAAAAAAAANNAAAAAA 1109

Qy        674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
              || |      ||| ||  |  ||   |  |    |    |     ||  ||  |    |
Db       1108 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAANNAAAAAAAAAAAAAAAAAAAA 1049

Qy        734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
              || |    |     ||   || |   |||    |    ||   |      |    |
Db       1048 AAAAAANAAAAANNAAANAAAAAAAAAAAAAANNAAAAAAANNAANAAAANAAANAANAN 989

Qy        794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
                | |     | |  |   ||     |   ||     |  ||||     | ||||| |
Db        988 AAANANNNNAANAANANAAAAAAAAAANNAAAAAAAAAAAAAAAANNAAAAAAAAANAAA 929

Qy        854 CTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACT 913
                  |   || |||  |||       |||||   |       | |    | |     |
Db        928 AAAAAAAAAAAAAAANAAAAAAAAAAAAAAAANNAAAAANAAAAAANNAANANAANAANA 869

Qy        914 TATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAAT 973
               |     || |  |        ||       | || |||    |||| ||       ||
Db        868 AAANNANAAAAANAAAAANNAAAAAAAAAAAAAAAAAAAANNNAAAAAAAAAANAAAAAA 809

Qy        974 TATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
               |   || ||| | |  |||||| | || ||  ||
Db        808 AAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAA 774


RESULT 4
CL082569
LOCUS       CL082569                1632 bp    DNA     linear   GSS 31-DEC-2003
DEFINITION  CH216-167P5_Sp6.1 CH216 Xenopus tropicalis genomic clone
            CH216-167P5, genomic survey sequence.
ACCESSION   CL082569
VERSION     CL082569.1  GI:40538482
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1632)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 1062
            High quality sequence stop: 1114.
FEATURES             Location/Qualifiers
     source          1..1632
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-167P5"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"
ORIGIN

  Query Match             8.7%;  Score 87.4;  DB 9;  Length 1632;
  Best Local Similarity   45.3%;  Pred. No. 2.6e-07;
  Matches  342;  Conservative    0;  Mismatches  409;  Indels    4;  Gaps    1;

Qy        254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
              | || ||||||||    | ||  |||  |  || | | || ||   |||| | |
Db        615 AAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAANATNAAAAAAAAAAAAAAAA 674

Qy        314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
               |  || |   ||  ||   || | | || |||  |  ||| | |||  | |   || ||
Db        675 AAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAATAAAAAAATAAAAAAAAAAAAAAATA 734

Qy        374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                    | || || | | | |||   ||||| |||     ||| |||| |      |  |
Db        735 AAAAAAAAAAAAAAAAAAAAAAAATAATTTAAAAAA----AAAAAAAAATAAAAAAAAAA 790

Qy        434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                  | | ||||      |   | |  | ||  |||| |         | | |||| ||
Db        791 AAATAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAATATAAAAAAAAAAA 850

Qy        494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
               ||||    ||   ||| || || ||   ||||||     | || || |  | || ||||
Db        851 AAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAA 910

Qy        554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
                 || || | || ||| ||| ||  |   | |||||  ||||  ||||      |  |
Db        911 AAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATATTAAAAAAAAAAA 970

Qy        614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
                   ||||    |          |   ||  ||     |        ||     ||
Db        971 AAAAAAAATATATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1030

Qy        674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
              ||  |  |   ||| ||  | |||   |  |    |  |||     ||  ||  |    |
Db       1031 AAAAAAATAAAAAAAAAAAAATAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1090

Qy        734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
              || |    |   | |||  || |   ||| |  |    || | |||  | |    |
Db       1091 AAAAAATAAAAAAAAAAAAAAAAAAAATAAAAAAAAAATAAATAATTAAAAAAAAA 1150

Qy        794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
                | |   | ||| |||  || |   ||  ||     |  || | || || ||||| |
Db       1151 AAAAATAAAAAAAAAAAAAAAAATAAAAAAAAAATAANANAAATAAAAATATAAAAAAAAA 1210

Qy        854 CTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACT 913
                 ||   || |||  |||       |||||  ||       | |    |       |
Db       1211 AAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTAAAAAATA 1270

Qy        914 TATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAAT 973
               |     || ||  |       ||||      | ||  ||    |||  ||  |     |
Db         1271 AAAAAATAAAATNAAAAAAAATAATAAAAAAAAAAATAAAAAAAAANAAAAATNANATAA 1330

Qy          974 TATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
                 |   || ||| | |  |||||| | || ||  ||
Db         1331 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1365


RESULT 5
CNS033GQ
LOCUS       CNS033GQ                1135 bp    DNA     linear   GSS 01-SEP-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
            208P24 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL226115
VERSION     AL226115.1  GI:7885026
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetradontoidea; Tetraodontidae; Tetraodon.
REFERENCE   1
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Estimate of human gene number provided by genome-wide analysis
            using Tetraodon nigroviridis DNA sequence
  JOURNAL   Nat. Genet. 25 (2), 235-238 (2000)
  MEDLINE   20296633
   PUBMED   10835645
REFERENCE   2
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
            Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
            Saurin,W., Bernot,A. and Weissenbach,J.
  TITLE     Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Genome Res. 10 (7), 939-949 (2000)
  MEDLINE   20359837
   PUBMED   10899143
REFERENCE   3  (bases 1 to 1135)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1..1135
                     /organism="Tetraodon nigroviridis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:99883"
                     /clone="208P24"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : COAG208DH12SP1-end :
                     PUC-Ori"
ORIGIN

  Query Match             8.6%;  Score 86.6;  DB 9;  Length 1135;
  Best Local Similarity   46.0%;  Pred. No. 3.8e-07;
  Matches  292;  Conservative    9;  Mismatches  330;  Indels    4;  Gaps    1;

Qy          254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
                | || ||||||||    |  |  |||  |  || | | ||||    |||| | |
Db          321 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 380

Qy          314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
                 |  || |   ||  ||   ||   | || |||  |  |||   |||  | |   ||  |
Db          381 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 440
```

```
Qy          374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                      |||| || | | | |||  ||   | ||| |   ||| ||||        |  |
Db          441 AAAAAAATAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAWAAAAA 500

Qy          434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                || ||   ||||      |   | |  | ||  |||| ||| |       |||| ||| |
Db          501 TTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTAAAAAAAAAAAAAAAATAT 560

Qy          494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
                 ||||    ||   ||| || || ||   || |||     | || || |  | || ||||
Db          561 TAAAA----TAAAAAAAAAAAAAAAAAAAAATAAAATAAAAAAAAAAAAAAAAAAWAAAA 616

Qy          554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
                || || || | |||||| ||| |   |   | | |||     |   |        | |
Db          617 TTAAAAAAAAAAAAAAAAAAAAATWAAAAAAAATAAAAATTTAAAAAAAATAAATATTTA 676

Qy          614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
                     |||  |  |:|  |     |   ||  ||    |||       ||     ||
Db          677 AAAAAAAAAAAWATWTAATAAAAAAAAAAAAAAAAATATATAAAAAAAAAAAAAAAAAAA 736

Qy          674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
                || ||  | | ||| ||  |  ||   |  |   :|  | |  :| ||  ||  ||  |
Db          737 AATTACATNTAAAAAAAAAAAAAAAAAAAAAMTCYAAAACAAAWTAAAAAAAAAATAATT 796

Qy          734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
                ||   : |:| | | |||  :|||   | ||  :|:   ||   ||     |    |
Db          797 AAMAYYATWAATAAAAAAAAWACAAAAACACACYAYAAAAAAAAAAAAAAAAAAAAAAAA 856

Qy          794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
                  | ||  | | | |||  || |   ||   |  :||| ||||| ||    ||||| | |
Db          857 AAAAAAAAAAANAAAAAAAAAATAAAAAAATAAAYTCAATAAAAAAAAACAAAAAAAAAC 916

Qy          854 CTGTATGGAAGAAATTAAACGCTGTGAAAAATCAA 888
                    |   ||  || |  |  |    |||||  ||
Db          917 AAAAAATAAAATAACTCTAAACCAAAAAAAAAAAA 951


RESULT 6
CNS0021J
LOCUS       CNS0021J                1101 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR05N11 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL061936
VERSION     AL061936.1  GI:4940214
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
```

```
     source          1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR05N11"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN

  Query Match             8.6%;  Score 86.4;  DB 9;  Length 1101;
  Best Local Similarity   39.7%;  Pred. No. 4.1e-07;
  Matches  329;  Conservative   44;  Mismatches  447;  Indels    9;  Gaps    1;

Qy        180 TGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGT 239
              |  | :| || :::  | ||  ||||||      || :   ||||:  | ||     : |
Db          6 TNNTTKTTTTTKKKTTTNTTNNTTTGTTTTNNTNTTTVNCTCTTTVTTTNTTNNNTTKTT 65

Qy        240 TGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGG 299
              |    ||     |   ||  |: |||||  || | ||  |   | ||:|  :: || |
Db         66 TTTTTTATTTTTTNNAAAANAWKAAAATNATAATTTTTTAATTTAAAKAAKWDTAAATAA 125

Qy        300 AAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAA 359
              ||| :| |     ||     || | |||      ||   | || |||  |  |||
Db        126 AAANHAAAAAAAAGTGTTTTCATGAAAGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAANN 185

Qy        360 ACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAA 419
               |       ||  |      | || || | | | |||   |   | ||| |   ||| ||
Db        186 TCNNNNCNNAANNAAAAAANAAAAAAAAAAAAAAAAAAACAAAAAAAAAAAAAAAAAAAA 245

Qy        420 AATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGC 479
              ||        |  |   ||   ||||      |   | |  | ||  |||| |   |
Db        246 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 305

Qy        480 TTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGA 539
                  :||| ||    ||||     |   ||| || || ||   ||||||     | || |
Db        306 AAAAMAAAAAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 365

Qy        540 ACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGT 599
              | |  | || ||||   || || | |||||| ||| ||  |   | |||||  ||||
Db        366 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 425

Qy        600 ATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGG 659
              |        |  |      |||  |              |  ||  ||     |
Db        426 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAANAAAAAAAAAAAAAAA 485

Qy        660 TCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTT 719
                ||     ||   ||  |      |||  |  ||| | |  |  : | | ||:   |||
Db        486 AAAAAAAAAAAAAAAAAAAANAANAAAAATATAATTTAWTTTTTTWTTAATTAWTTTTTT 545

Qy        720 AAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAA 779
                 |   | | ||:       | | ||  ||:         |:|  ||   | || : :
Db        546 TTTTTTTTTTTTTWTTAATTTTAATTTTTAAW---------AWAAATTTAATAAAAWAWT 596

Qy        780 TCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGA 839
              :  | :   ||  :| |::|  | |||      ||  : :        :| | | ||  :
Db        597 WTTTAWTTTTAATWTAAWWAAAAAAAAWTTTTAAAWAWTWTTTTTTTTWTTAATTAATAW 656

Qy        840 ACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATAT 899
              |  | |:| | |   | | |  ::  :|   :|   | | ||||:|       ||  | |
Db        657 AAAAAAWATAWAAWTTWTTTATWWATWAAATWAAATTTTAAAAAWTTTTTTAWTTTTTTT 716

Qy        900 TTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAA 959
              :||  |   |   |||:    || |:    :  :| ||||| || :  |  ||:   |:|
Db        717 WTTTAAAAWTTTATTAWTTTTAAAAWATWAWTWWTAAATTTTTTAWATAWTAAWATTAWA 776

Qy        960 AGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
              | |  || |    | || |     |  ||  |||||| | || || :||
Db        777 AAATTTTTTAATTTAATTTTTTTTAATTAAAAAAAAAAAAAAAAAAWAA 825


RESULT 7
BH651171/c
LOCUS       BH651171                 773 bp    DNA     linear   GSS 19-FEB-2002
DEFINITION  BOMJP20TR BO_2_3_KB Brassica oleracea genomic clone BOMJP20,
            genomic survey sequence.
ACCESSION   BH651171
VERSION     BH651171.1  GI:18709113
KEYWORDS    GSS.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 773)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1..773
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOMJP20"
                     /clone_lib="BO_2_3_KB"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
ORIGIN

  Query Match             8.3%;  Score 83.6;  DB 8;  Length 773;
  Best Local Similarity   51.0%;  Pred. No. 1.5e-06;
  Matches  197;  Conservative    0;  Mismatches  189;  Indels    0;  Gaps    0;

Qy        254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
              |||| || ||||||   | ||  ||   |   | | |||||| |  |||  | | |
Db        744 ATAAAAATAAAAATAAAAATAAAAATAAAAATAAAAATAAAAATAAAAATAAAAATAAAA 685

Qy        314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
                  ||     ||  ||  |||     || | |  | ||||  ||||    |    |  |
Db        684 ATAAAAATAAAAATAAAAATAAAAATAAAAATAAAAATAAAAATAAAAATAAAAATAAAA 625

Qy        374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
               |    |||| |||| | | |||  ||   ||||| |   |||||||| |     ||  |
Db        624 ATAAAAATAAAAATAAAAATAAAAATAAAAATAAAAATAAAAATAAAAATAAAAATAAAA 565

Qy        434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
               | ||   |||| |    ||  | || | ||| ||||||    |     |||||  |
Db        564 ATAAAAATAAAAATAAAAATAAAAATAAAAATAAAAATAAAATTAAAAATAAAAATAAAA 505

Qy        494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
                |||  |  |  |||| |  || |    || |||     | |  || |  | || ||||
Db        504 ATAAAAATAAAAATAAAAATAAAAATAAAAATAAAAATAAAAATAAAAATAAAAATAAAA 445

Qy        554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
               | || |||| |||||| ||| |   |   | ||| |  || ||  |  |     |  ||
Db        444 ATAAAAAATAAAAAAAATAAAAATAAAAATAAAAATAGAAATAGAGAAGGGAATGATAAG 385

Qy        614 CTGCTAAATCAGGTTTGCTTGCTCAT 639
               |  | | |  | ||  | || || |
Db        384 ATAATGACTATGCTTCACCTGTTCTT 359


RESULT 8
CR734084/c
LOCUS       CR734084                1172 bp    mRNA    linear   HTC 19-AUG-2004
DEFINITION  Tetraodon nigroviridis full-length cDNA.
ACCESSION   CR734084
```

```
VERSION     CR734084.1  GI:51232422
KEYWORDS    HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE      Tetraodon nigroviridis
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetradontoidea; Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 1172)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
            : 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     The sequences are based on single pass reads.
            More information available at
            http://www.genoscope.cns.fr/tetraodon.
FEATURES             Location/Qualifiers
     source          1..1172
                     /organism="Tetraodon nigroviridis"
                     /mol_type="mRNA"
                     /db_xref="taxon:99883"
                     /tissue_type="fish"
ORIGIN

  Query Match             8.3%;  Score 83.6;  DB 3;  Length 1172;
  Best Local Similarity   47.2%;  Pred. No. 1.5e-06;
  Matches  323;  Conservative    0;  Mismatches  354;  Indels    7;  Gaps    2;

Qy        210 TGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCG 269
              | | | | || ||   || | |||  | ||   ||     |  | || ||||||||
Db        852 TCCTTAACTTTTTATTATTTAAAAATTTTTAAAATTAATTAAAAAAAAAAAAAAAAAAAT 793

Qy        270 TATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCA 329
              | |||  |||  |  ||     |||| ||  |||| | | ||          || |  ||  |
Db        792 TTTTAAAAAAAAAAAAAAATTTTAAATTTTTAAAAAAAAATTTAAAATTAAAAAAAATAA 733

Qy        330 ACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAATAG 389
              |   || ||  || ||||    ||| |||||     || |||||        | || || |
Db        732 AAAAAAATTTAAAAAAAATTTAAAAAATTTAAAAAATTTTAAATTAAAAAAAAAAAAAAA 673

Qy        390 ACACAAAGGTATTTAT-----AAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAA 444
              |   |||  |||| |     ||| |   ||| ||||        |  |   ||   |||
Db        672 ATTAAAAAAAATTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 613

Qy        445 AGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTA 504
              | |||  |     |  | ||  ||| ||     |     | ||| || ||  |||     |
Db        612 AATTTTAAATTTTAAAAAAAAAAAAATTAAATTTAAAAATTAAAAAAATTTAAAAAAATT 553

Qy        505 GGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATAT 564
                ||   || || ||   ||||||    |  || |||    | || ||||| || ||||
Db        552 TTTATTAAAAAAAAAAAAAAAAAAAAAAATTAAATAACTTAAAAAAAAAATCTAAAATAA 493

Qy        565 AAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCA 624
              |||||| ||| |   |   | ||||||  ||  |||          | ||      |||  |
Db        492 AAAAAAAAAAGACTAAAAAAAAAAAAAT--AATAGTAAAAAAAAAAAATAAAAAAAAAAA 435

Qy        625 GGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTT 684
                     |     |    | ||     |        ||      ||    ||  |   |||
Db        434 AAAAAAATAAAATAAATTAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAATTT 375

Qy        685 AAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTAC 744
              ||| ||  |  ||   |  ||   |  |||     ||   | ||   ||||  |  |  |
Db        374 AAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAATAATAAAATAAAAATTAAT 315

Qy        745 TTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACA 804
              | | |||  || |   |||  ||| | ||   | |    |   ||      | ||  | |
Db        314 TAAAAAAAAAAAAAAAAAAAAAATATTTAAAAAAAATTAAATATAATAAAAAAAAAAAAAA 255

Qy        805 AACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAG 864
              || ||   ||     ||  ||      | ||||| ||  |  |||| |   |  |   ||
Db        254 AAAAATAAAAAAAAAAAAAAAAAAAAATTAAAAAAAAAAAATAAAAAAAAATAAAAAAAT 195
```

```
Qy        865 AAATTAAACGCTGTGAAAAATCAA 888
              |||  ||    | | |||||  ||
Db        194 AAAAAAATTTATATAAAAAAAAAA 171

RESULT 9
AG429915/c
LOCUS       AG429915                1362 bp    DNA     linear   GSS 03-JUN-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-304L12.TJ, genomic survey
            sequence.
ACCESSION   AG429915
VERSION     AG429915.1  GI:48072978
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus
  ORGANISM  Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     BAC end Sequences of Library MSMg01
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1362)
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            Tsukuba Institude, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : TJ
            LIBRARY
            Vector     : pBACe3.6
            R.Site 1   : EcoRI
            R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
     source          1..1362
                     /organism="Mus musculus molossinus"
                     /mol_type="genomic DNA"
                     /sub_species="molossinus"
                     /db_xref="taxon:57486"
                     /clone="MSMg01-304L12.TJ"
                     /sex="male"
                     /tissue_type="mixture of kidney and spleen"
                     /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN

  Query Match             8.2%;  Score 82.4;  DB 9;  Length 1362;
  Best Local Similarity   36.2%;  Pred. No. 2.6e-06;
  Matches  272;  Conservative    0;  Mismatches  480;  Indels    0;  Gaps    0;

Qy        257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
              | ||||| |||    |  |  |||      || | | ||||    |||| | |       |
Db       1331 ANAAAANAAAAAAANAAAAAAAAAAAANNAAANAAAAAAAAAAANNNAAAAAAAAAANNNNNNAA 1272

Qy        317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
               || |   ||  ||    |   | || |||       ||    |||  | |    || |
Db       1271 AAAAAAAAAAAAAAAAAANNAAAAAAAAAAAAAAANANAANAANNNAAAAAAAAAAAAAAANAAAAA 1212

Qy        377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
                   || || | | | |||    |      ||| |   | |  |||           |
Db       1211 ANNNNAAAAANAAAAAAAAAAANAAAANNAAAAAAAAAAANAANAAAAAANNANNANAAAANA 1152

Qy        437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
              ||    |||       |     |  | ||  |||| |    |       |||| ||    ||
Db       1151 AAAAAAAANNAAAANAAAANNAAAAAAAAAAAAAAANANAAANANAAANAAAAAAAAAAAAA 1092
```

```
Qy         497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
               ||     |   ||| || || ||     ||||     | |  || |  | ||   |
Db        1091 AANAAAAAANNAAANAAAAAAAANNANNAAAAAAAAAAAANNAAAAAAAAAAANNANAAA 1032

Qy         557 AAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTG 616
               || || | |||| | ||| ||  |      ||      ||            |  |
Db        1031 AAAAAAAAAAAANAAAAAAAAAAAAANNNNAANNAANNAANNNNAANANNAAANAAAAAA 972

Qy         617 CTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAAC 676
                 |||  |             |   ||         |        ||     ||
Db         971 AAAAAAAANANNNAAAAAANNAAAAAAANNNANAAAAAAAAAAAAAAAAAAAAAAAANNN 912

Qy         677 TAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAG 736
                |      ||| ||  |  ||   |  |    |  |||         ||        ||
Db         911 AAAAAAAAAAAAAAAAAAAAAAAAANNAAAAAAAAAAANAANANNANAAAANNNAANAAA 852

Qy         737 GACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCA 796
                |    |   | |||  || |   |||    |    ||   ||     |    |     |
Db         851 AANNNNAANAAAAAAAAAAAAAAAAAAAANNAAAAAAAAAAAANNAAAAAANAANNAAAA 792

Qy         797 TAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTG 856
                ||  | ||| |||  ||     ||  ||     |  ||   ||  | ||| | |
Db         791 AAAANANAAANAAAANAAAAAAAAAAAAAAAAANAAAAANNAAAAAAAAAANAAAAAAAA 732

Qy         857 TATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTAT 916
                |   || |||  ||        |||||  ||       | |    |       |   |
Db         731 AANAAAAAAAAANAANNAANAAAAAAAAAAAAAAAAAAAAAAANANANNAAAAAAAAAAA 672

Qy         917 GGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTAT 976
                   || |  |         |       |  | |||    ||||  |       ||  |
Db         671 NANAAAAAAAAAAANNAAANAAAAAAANAANAAAAAANNAAAAANNAANAAAAAAANAAA 612

Qy         977 CTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
                  |  ||   |  ||     | || ||  ||
Db         611 AANAANAANNNAAAAANNNNAAAAANAANNAA 580

RESULT 10
CL461494
LOCUS       CL461494                1042 bp    DNA     linear   GSS 31-MAR-2004
DEFINITION  SAIL_1148_D03.v1 SAIL Collection Arabidopsis thaliana genomic clone
            SAIL_1148_D03.v1, genomic survey sequence.
ACCESSION   CL461494
VERSION     CL461494.1  GI:45864399
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1042)
  AUTHORS   Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D.,
            Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D.,
            Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kimmerly,B.,
            Mitzel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.
  TITLE     A high-throughput Arabidopsis reverse genetics system
  JOURNAL   Plant Cell 14 (12), 2985-2994 (2002)
  MEDLINE   22356987
   PUBMED   12468722
COMMENT     Contact: Sessions A
            Applied Trait Genetics
            Syngenta Biotechnology Inc.
            3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
            Email: allen.sessions@syngenta.com
            ABRC Stock Number CS842250; T-DNA left border flanking sequences of
            Syngenta Arabidopsis Insertion Library (SAIL) lines are available
            through the Arabidopsis Biological Resource Center (ABRC).
            Sequences represent a pool of amplified genomic regions and not
            single contiguous sequences.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1..1042
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /clone="SAIL_1148_D03.v1"
                     /clone_lib="SAIL Collection"
                     /note="T-DNA left border sequences were isolated using a
                     modified TAIL-PCR strategy"
ORIGIN

  Query Match              8.1%;  Score 81.6;  DB 9;  Length 1042;
  Best Local Similarity   54.1%;  Pred. No. 3.8e-06;
  Matches  185;  Conservative    0;  Mismatches  153;  Indels    4;  Gaps    1;

Qy         254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
               || | ||||||||   || || ||||  |   | | ||||||    |||||| | ||
Db         601 ATTAAAAAAAAAAATATAATAATAAAAAAAATAAAAATAAAAAAAAAAAATAAATTTTAA 660

Qy         314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
               ||  || |   ||   |   ||   | |  ||   |   |    ||   | || |||  |
Db         661 TAATAAAAATAAAAATATAAAATAAAAATTAATAAAAATATAAAAATATAAATATAAAAA 720

Qy         374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
               |   | | || || | | | |||   |  || ||| |  |||||||||    |   |  |
Db         721 TAAATAAAAAAAAAAAAAAAAAAATANAAATAAAAANATATAAATAAAAAAAATTTAAAAA 780

Qy         434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
               || ||   ||||  |   |   | |  | ||| | |||||| |   | | |||  |  |
Db         781 TTAAAAAAAAAAAATAATAATAATAAAAAAATNATAATATTAAAAATATAAAATAATATA 840

Qy         494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
                |||    | |   ||| || || ||  | |||||   | | || || |  |||| |
Db         841 AAAA----TAATAAAAAAAAAAAAAAATTTAAAAATAATAAAAAAAAAAAAATAATACTG 896

Qy         554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAA 595
                  || ||||||||||| |||  |  ||  |||||||  | |
Db         897 AAAAAAAATATAAAAAAAAAAANATAATAAATAAAAAATATA 938

RESULT 11
CL101048
LOCUS       CL101048                1017 bp    DNA     linear   GSS 05-JAN-2004
DEFINITION  ISB1-3604_Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-3604,
            genomic survey sequence.
ACCESSION   CL101048
VERSION     CL101048.1  GI:40594683
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1017)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 75000   Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 13
            High quality sequence stop: 136.
FEATURES             Location/Qualifiers
     source          1..1017
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:8364"
                     /clone="ISB1-3604"
                     /clone_lib="ISB1"
```

```
                    /note="Vector: pBeloBAC11; ISB-1 Xenopus tropicalis BAC
                    Library Segment 1"
ORIGIN

  Query Match              8.1%;  Score 81.4;  DB 9;  Length 1017;
  Best Local Similarity   54.2%;  Pred. No. 4.1e-06;
  Matches  188;  Conservative    0;  Mismatches  156;  Indels    3;  Gaps    1;

Qy          254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
                | |  ||||| ||    | ||  |||  |  ||   |||||| |  | || | | |
Db          631 AAATAAAAAATAAAAAAAATAAAAAAATAAAAAATAATAAAAATAAATAAAAAAATAAAA 690

Qy          314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
                 |  || |   ||  ||  ||   || |  |||| |   ||  || |  | || |||| |
Db          691 AAAAAATATAAAATTAAAATATAATAAATAAAATTATATAATATATATAATATATAATAA 750

Qy          374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                 |  | | |  |  |   | |||   ||  | ||  |   | | |||| |  |  ||  |
Db          751 ATAATAAAATAATAAATAAAAAATTAATAAAAAATAAAAAATAAAAAAATAATATTAAAA 810

Qy          434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                |  ||   |||| |||  |   | | |||||   ||| |   ||    |  ||| ||  |
Db          811 TAAAATAAAAAAATTTTAA---ATAATATAAAATAAAAAAAAATATAATATAAATAAAAT 867

Qy          494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
                 ||||    ||  |||| || || ||  |||||||     | || || |  |||| ||||
Db          868 AAAAAAAAATAAATAAATAAAAATAAATTAAAAAAAATAAAAAATAAAATAATAAAAAAA 927

Qy          554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTA 600
                 | || |||| | |||| ||||||  ||  ||||||   | ||  ||
Db          928 ATAAATAATAAATAAAAAAAATAAAAATAAATAAAATAAATAAAATA 974


RESULT 12
CF238805
LOCUS       CF238805                1626 bp    mRNA    linear   EST 05-AUG-2003
DEFINITION  AGENCOURT_15099447 NICHD_XGC_Emb6 Xenopus tropicalis cDNA clone
            IMAGE:6995950 5', mRNA sequence.
ACCESSION   CF238805
VERSION     CF238805.1  GI:33442013
KEYWORDS    EST.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1626)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Robert M. Grainger
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14680  row: g  column: 21
            High quality sequence start: 71
            High quality sequence stop: 316.
FEATURES             Location/Qualifiers
     source          1..1626
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="IMAGE:6995950"
                     /tissue_type="neurula"
                     /dev_stage="embryo, stages 14-19"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NICHD_XGC_Emb6"
                     /note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV;
                     Cloned unidirectionally. Primer: Oligo dT. Average insert
                     size 2.1 kb. Constructed by Invitrogen. Note: This is a
                     Xenopus Gene Collection (XGC) library."
ORIGIN

  Query Match              8.1%;  Score 81.4;  DB 7;  Length 1626;
  Best Local Similarity   44.9%;  Pred. No. 4.2e-06;
  Matches  340;  Conservative    0;  Mismatches  413;  Indels    4;  Gaps    2;

Qy          256 AACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTA 315
                || ||||  ||   ||  | ||||  |  || | | ||||    |||| | | |     |
Db          866 AAAAAAATTAAATTTAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAATATTAAAAAA 925

Qy          316 GGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATT 375
                  || |   ||  ||   ||   | || |||| |  |||   |||  | | | ||  |
Db          926 AAAAAAAAAAAAAAAAAAAANAAAAAAAAAATAAAAAAAAAAAAAAAAAAATAAAAAAAA 985

Qy          376 GCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATT 435
                  | |||  || | | | |||   |   | ||| | | ||| ||||        |  |
Db          986 AATAATATAAAAAAAAAAAAATAAAAAAAAAAAAAATNAAAAAAAAAAATAAAAAAAAAA 1045

Qy          436 GAACTGAAAAGTTTCGATGGAGATTATAATGAAAAT---ATTGATGCTTTTAAAACAATT 492
                 ||   ||||      ||  | |  | ||  ||||    ||| |       |||| ||
Db         1046 AAAAAAAAAAAAAAAAATAAAAAAAAAAAATAAAAAAAAATTAAAAAAAAAAAAAAAAAA 1105

Qy          493 TCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAA 552
                  ||||    ||   ||| ||    ||   ||||||     | || || |  | || |||
Db         1106 AAAAAAAAAATAAAAAAAAAAATTAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAA 1165

Qy          553 ATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTA 612
                |   ||  | |  ||||| ||| ||  |   | |||||  ||||     |   |  | ||
Db         1166 AAAAAAATAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATATATATATAATA 1225

Qy          613 GCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGT 672
                      |||  |             ||  ||  ||     |  |     ||    |||
Db         1226 AAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAATATAAAAAAAAATAAAA 1285

Qy          673 CAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTT 732
                 ||  |  |   ||| ||  |  |  |    |   ||  |||    |||  ||  |
Db         1286 AAAAAAAATAAAAAAAAAAAAAAANTTTNNTAAAATAAAAAAAAAATAATAAAAAAAAAA 1345

Qy          733 AAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATG 792
                ||   |  |     | |||  || |   |||    |    || | ||     |   ||
Db         1346 AANTNAANTNNTNAAAAAAAAAAAAAAAAAATAAAAAAAAAAATAAAATTNNAAAATAAA 1405

Qy          793 TTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGA-AAAAGA 851
                     ||  | ||| | |  || ||  ||  ||  || |  |    ||  | |  ||| |
Db         1406 AAATAAAATATAAATATATAAAATTAAAAAAAAAATTAAGAATTTAAATNAAAGGAAAAA 1465

Qy          852 TCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGA 911
                   |  |   || |||  |||   | | |||||| ||    |  | |    | |     |
Db         1466 ATNTAAAATAAAAAAAAAAAAAAATATAAAAAATAAAATAATAAAAAAAAAAAAAATAAA 1525

Qy          912 CTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGA 971
                   ||        | |  |   |||||   |  | |  |||     ||   |       |
Db         1526 AANATNTNATNATTAAAATAAATTAATAAATAAAAATAAAANTAATAATATATAAAAAAA 1585

Qy          972 ATTATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
                ||||| ||| ||| |    |||||| |  | ||  ||
Db         1586 ATTATATAAAAAAAAATTNAAAAAAAAANAAAAAAAA 1622


RESULT 13
CL038406
LOCUS       CL038406                1594 bp    DNA     linear   GSS 31-DEC-2003
DEFINITION  CH216-46A6_Sp6.1 CH216 Xenopus tropicalis genomic clone CH216-46A6,
            genomic survey sequence.
ACCESSION   CL038406
VERSION     CL038406.1  GI:40494319
KEYWORDS    GSS.
```

```
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1594)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 471
            High quality sequence stop: 522.
FEATURES             Location/Qualifiers
     source          1. .1594
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-46A6"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"
ORIGIN

  Query Match             8.1%;  Score 81.2;  DB 9;  Length 1594;
  Best Local Similarity   44.0%;  Pred. No. 4.6e-06;
  Matches  332;  Conservative    0;  Mismatches  423;  Indels    0;  Gaps    0;

Qy         254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
               | || ||||||||       | ||||  |  || | | |||| |  |||| | |
Db         756 AAAAAAAAAAAAAAAAATAAAATAAAAAAAAAAAATAAAAAAATAAAAAAAAAAAAAAAA 815

Qy         314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
                |  || |   ||  ||   ||   | || | |  |  |||  ||||  | ||  ||  |
Db         816 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAATAAAAAAAATAAAAAAA 875

Qy         374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                     | || || | | | |||   |||   ||| |   ||| ||||        |  |
Db         876 AAAAAAAAAAAAAAAAAAAAAAAAAAATTATAAAATAAAAAAAAAAAAAAAAATAAAAAA 935

Qy         434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                  ||   ||||      |   | |  | ||| |||| ||  |       |||| ||
Db         936 AAAAAAAAAAAAAAAAAAAATAAAAAAAAAATAAAAAAATATAAAAAAAAAAAAAAAAAA 995

Qy         494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
                ||||     |   ||| || || ||   |||||      | || || |  | || | ||
Db         996 AAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAATAA 1055

Qy         554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
                  || |||| |||||  ||| ||  |   | |||||  | ||   ||       |  |
Db        1056 AAAAAAAATAAAAAAATAAAAAAAAAAAAAAAAAAAAAAANAAATAATAAAAAAAATAAA 1115

Qy         614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
                   ||||           |  | |   ||  ||     |  |     |  |   ||
Db        1116 AAAATAAAAATAAAAAAAATAATAAAAAAAATAAAAAAAAAATATAAAATATAAAAAAAA 1175

Qy         674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
               ||         ||| ||  |  ||      |    | ||||   | ||  ||  |    |
Db        1176 AAAATAAAAAAAAAAAAATAAAAATAATAAAAAAAAATAAAAAATAAAAAAAAAAAAAAA 1235

Qy         734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
               ||  |    |   | |||| || |   |||    |  |  |   ||     |    ||
Db        1236 AAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAATATAAAAAAAAAAAAAAAAATTA 1295
```

```
Qy         794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
                 | ||  | ||| |    ||     ||  |   |  |  |||| ||  | ||||| ||
Db        1296 AAAAAAAAAAAAAAATNAAAAAAATAAAAAANAATAAAAAAAAAAAAAAAAAAAAAAATA 1355

Qy         854 CTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACT 913
                   ||  || |||  |||       || ||   |       | |    | |     |
Db        1356 AAAAATAAAAAAAAAAAAAATAATAAAANAAAATAATATAAAAAAAAAAATAAATAAAAA 1415

Qy         914 TATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAAT 973
                |     || |  |  |     |||   |  | || |||    | || ||       ||
Db        1416 AAAATAAAAAAATAAATAAAAAAATANATAAAAAAAAAAAAAAATAATAAAAAAAAAAAA 1475

Qy         974 TATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
               ||   || ||| | |  |||||| | || ||  ||
Db        1476 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1510


RESULT 14
CNS00HX9/c
LOCUS       CNS00HX9                1101 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC:
            BACR35M04 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL073856
VERSION     AL073856.1  GI:4953838
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR35M04"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN

  Query Match             8.0%;  Score 80:8;  DB 9;  Length 1101;
  Best Local Similarity   44.7%;  Pred. No. 5.5e-06;
  Matches  169;  Conservative   31;  Mismatches  178;  Indels    0;  Gaps    0;

Qy         232 AAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGAT 291
               ||| |||      ||    |  |:|| ||||||||       |   ||  | :|| : |
Db         449 AAAACTGAAAATGTAAGAAAAAAWAAAAAAAAAAAAAAATAAAAATAAAAAAWAAAWAAA 390

Qy         292 AAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGT 351
               ||||:   |||| | :    |  :  || |   ||  ||   || : | || |||  |
Db         389 AAAAWAAAAAAAAAAWAAAATAAWAAAAAAAAAAAAWAAAAAAAGWAAAAAAAAAAAAAA 330
```

```
Qy        352 AAACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGAC 411
              ||| : |||  | |   ||  |:     : || |||| | | :|| : |  :::|:: |
Db        329 AAAAWAAAAAAAAAAAAAAAAAWAAWAAWAAAAAATAAAAAAWAAARAAAAWWWAWMAAA 270

Qy        412 TTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAAT 471
              : ||| ||||     |  |  | : ||   ||||      |   | :    ||  ||||
Db        269 WAAAAAAAAAAATAACAAAAWAAWAAATAAAAAAAAAAAWAAAAAAWAATAAAAWAAAAA 210

Qy        472 ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGC 531
              |   |   :|  |||| ||  : :||| |||:|   ||| || || ||    |||||| |
Db        209 AAACAATTWTAAAAAAAAAMAWAWAAAACTTWAAWAAAAAAAAAAAAAWWATAAAAACCC 150

Qy        532 TTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAAT 591
                | |  || |  | :   |||   || || : || |||  ||:||: |   | ||:||
Db        149 AGAAATTAATAGWAAWGAGAAAAAAAATAAAWAAATAAAATAAWAAMAAAAAAAAAWAAA 90

Qy        592 CAAAAGGTATTGCCTGCA 609
               :||:  || |   | ||
Db         89 AWAAWAWTAATAATTTCA 72
```

```
RESULT 15
AG386893/c
LOCUS       AG386893                1519 bp    DNA     linear   GSS 03-JUN-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-201E07.TJ, genomic survey
            sequence.
ACCESSION   AG386893
VERSION     AG386893.1  GI:47998098
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus
  ORGANISM  Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     BAC end Sequences of Library MSMg01
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1519)
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            Tsukuba Institute, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : TJ
            LIBRARY
            Vector     : pBACe3.6
            R.Site 1   : EcoRI
            R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
     source          1..1519
                     /organism="Mus musculus molossinus"
                     /mol_type="genomic DNA"
                     /sub_species="molossinus"
                     /db_xref="taxon:57486"
                     /clone="MSMg01-201E07.TJ"
                     /sex="male"
                     /tissue_type="mixture of kidney and spleen"
                     /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN

  Query Match             8.0%;  Score 80.8;  DB 9;  Length 1519;
  Best Local Similarity   41.9%;  Pred. No. 5.5e-06;
```

```
  Matches  316;  Conservative    0;  Mismatches  439;  Indels    0;  Gaps    0;

Qy        254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
              | || ||||||||    |  |  |||  |  || | |  |||    |||| | |
Db       1108 AAAAAAAAAAAAAAAAAANAAANAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAANN 1049

Qy        314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
               |  || |    |  ||   ||   | || |||  |  |||   ||   | |   ||  |
Db       1048 AAAAAANAANANAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAANAAAAAANAAAAAA 989

Qy        374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                    | || || | | | |||   |   | ||| |   ||| ||||        |  |
Db        988 ANNAANAAAAAAAAAAAAANAAAAAAANNAANAAAAAAAAAAAAAAAAAAAAAAAAAAAA 929

Qy        434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                 ||   ||||      |   | |  | ||  |||| |   |       |||| ||
Db        928 ANNAAAAAAAAAAAANAAAAAAAAAAAANAAAAAAAAAAAAAAAANNAAAAAAAAAAAAA 869

Qy        494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
               ||||         ||| || || ||   | |||| |   | || || |  | || ||||
Db        868 AAAAANAAAANNAAAAAAAAAAAAAAAAAANAAAAAGAAAAAAAAAAAANAAAAANAAAA 809

Qy        554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
                 || || | ||||||   | ||  |  || |||||  |||    |        |  |
Db        808 AAAAAAAAAAAAAAAAAANNAAAAAAAAAGAAAAAAAAAAAANNAAAAAANAAAAAAAAA 749

Qy        614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
                   |||                |   ||  ||     |        ||     ||
Db        748 AAAAAAAAAANNAAAAANNAAAAAAANAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAA 689

Qy        674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
              ||  |      ||| ||  |  ||   |  |    |  |||     ||  ||  |    |
Db        688 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 629

Qy        734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
              ||  |    |   | |||  || |   |||    |    ||   ||     |    |
Db        628 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAANAAAAAA 569

Qy        794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
                | ||  | ||| |||  ||     ||  ||     |  |||| ||  | ||||| |
Db        568 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 509

Qy        854 CTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACT 913
                  |   || |||  |||       |||||  ||       | |    | |     |
Db        508 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAA 449

Qy        914 TATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAAT 973
               |     || |  |        ||       | || |||    |||| ||       ||
Db        448 AAAAAAAAAAAAAAAAAAAAAAAANAAAANAAAAAAAAAAAAAAAAAAAAAAAAANAAAA 389

Qy        974 TATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
               |   || ||| | |  |||||| | || ||  ||
Db        388 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 354
```

```
RESULT 16
CL078552
LOCUS       CL078552                1402 bp    DNA     linear   GSS 31-DEC-2003
DEFINITION  CH216-151D21_Sp5.1 CH216 Xenopus tropicalis genomic clone
            CH216-151D21, genomic survey sequence.
ACCESSION   CL078552
VERSION     CL078552.1  GI:40534465
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1402)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
```

```
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: Sp5 atctgccgtttcgatcct
            Class: BAC ends
            High quality sequence start: 867
            High quality sequence stop: 918.
FEATURES             Location/Qualifiers
     source          1..1402
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-151D21"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"
ORIGIN

  Query Match             8.0%;  Score 80.6;  DB 9;  Length 1402;
  Best Local Similarity   40.3%;  Pred. No. 6.1e-06;
  Matches  203;  Conservative    0;  Mismatches  301;  Indels    0;  Gaps    0;

Qy          93 AAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTAC 152
               ||  ||   ||  ||| ||     |||| |          |||||  |  ||||||
Db         141 AAAAAAAAAAAAGAAAAAGANGAAAAAANANANNNACACNATCAAAAAAAAAGAAGAGNN 200

Qy         153 TACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGC 212
                                                    ||| |  |   |      |
Db         201 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNANNNNNNCCTTCAAAAAAAAAAAAAAGNNNN 260

Qy         213 GTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTAT 272
                                    |           |    |  | || ||||||||    |
Db         261 NNNNNNNNNNNNNNNNNNNNNANNNNNANANNNANAAAANAAAAAAAAAAAAAAAAAAAN 320

Qy         273 TATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACC 332
                |  |||  |  || | | ||||    |||| |||       |  || |   ||  ||
Db         321 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAAAAAAAAAAAAAGAAAAAANAAAAA 380

Qy         333 TAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAATAGACA 392
                ||   | || |||  |  ||    |||  | |   ||  |      | || || | | |
Db         381 AAAAAAANAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAA 440

Qy         393 CAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGA 452
                |||   |   | ||| |   ||| ||||        |  |   |    ||||      |
Db         441 AAAAAAAAAANAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAANAAAAAAAAAAAAAA 500

Qy         453 TGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGA 512
                  | |  | ||  |||| |   |       |||| ||    ||||     |   ||| |
Db         501 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 560

Qy         513 AGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGA 572
               | ||  |   | ||||     | || || |  | |||||||   || || | |||||  |
Db         561 AAAANNAAAAANAAAAAAANAAAAAAAAAAAAAAAAGAAAANAAAAAAAAAAAAAAANAA 620

Qy         573 AATAACTATGGATAAAAATCAAAA 596
               || ||  |   | || ||  ||||
Db         621 AAAAAAAAAAAAAAANAAAAAAAA 644

RESULT 17
AG278469/c
LOCUS       AG278469                1506 bp    DNA     linear   GSS 02-JUN-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-048M18.T7, genomic survey
            sequence.
ACCESSION   AG278469
VERSION     AG278469.1  GI:47851346
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus
  ORGANISM  Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     BAC end Sequences of Library MSMg01
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1506)
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            Tsukuba Institude, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : T7
            LIBRARY
            Vector     : pBACe3.6
            R.Site 1   : EcoRI
            R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
     source          1..1506
                     /organism="Mus musculus molossinus"
                     /mol_type="genomic DNA"
                     /sub_species="molossinus"
                     /db_xref="taxon:57486"
                     /clone="MSMg01-048M18.T7"
                     /sex="male"
                     /tissue_type="mixture of kidney and spleen"
                     /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN

  Query Match             8.0%;  Score 80.4;  DB 9;  Length 1506;
  Best Local Similarity   46.6%;  Pred. No. 6.7e-06;
  Matches  290;  Conservative    0;  Mismatches  331;  Indels    1;  Gaps    1;

Qy         251 ATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTT 310
               ||  ||  ||||||||    | ||  |||  |  || | |||||| |  |   || | |
Db         875 ATAGTATTAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAATAAAAATCAAGGGTAAAATA 816

Qy         311 CTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAA 370
                   |  || |   |   ||    |  || || ||   | || |   |||  | |   ||
Db         815 AAAAAAAAAAAATAATAAAAATAGAAATAAAAAAATAAAATATATAAAAATAAAAAAAAA 756

Qy         371 TTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTA 430
                ||      |||  || ||| | |||   || || ||| |  |||||||||       |
Db         755 ATAAAAAAAATATAAAAAGAAAAAAAAAAATATAAAAAAAAATAAATAAAAAAAAAAATT 696

Qy         431 CGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAA 490
                     ||   |||| | |  |   | | ||  || |||| | | |       |||| ||
Db         695 AATAAAAATAAAAAAATATAAAATAAAAATAATATAAAAAAAATAAAAAAAAAAAAAAAA 636

Qy         491 TTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGA 550
                   || |  | ||   ||  || || ||   ||||||     | || || |  | || |
Db         635 AAAAAAGAAATATAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAAAA 576

Qy         551 AAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAG 610
               |||   || || | |||||| ||| ||  |   | |||||  ||||   |        |
Db         575 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATATAATAA 516

Qy         611 TAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAA 670
                |      |||| |  |          |   ||  ||    ||      | ||     |
Db         515 AAAAAAAAAAATAAAATAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAATAAAGGGGGAG 456

Qy         671 GTCAACTAGGTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATC 730
```

```
                 |  ||      ||| ||    |   || || | ||| ||||     || ||  ||
Db        455 AAAATATAAAAAAAAAAAAATATATTGTTGTTTTTATTAAAAAAAAAAAAATTATTTAAA 396

Qy        731 TTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTA 790
                |||  |    |  || |||   | |   |||    |  | || | ||     |    |
Db        395 AAAAAAAAAAAAAAATAAAAAAATAAAAAAAAAAAGAAAATAAAAT-AAAAAAAATAAAA 337

Qy        791 TGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAG 850
                     ||  | ||| |||| ||     ||  ||     |  || |  || | |||||
Db        336 AAAAATAAAAAAAAAAAAAAGAAAAAAAAAATAAAAAAAAAAAAACAATACAAAAAAAAC 277

Qy        851 ATCCTGTATGGAAGAAATTAAA 872
              |      ||   | ||||||||
Db        276 AAAACAAATAATAAAAATTAAA 255


RESULT 18
CG745927/c
LOCUS       CG745927                1225 bp    DNA     linear   GSS 24-OCT-2003
DEFINITION  P039-1-A06.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG745927
VERSION     CG745927.1  GI:37966853
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 1225)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
  TITLE     An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
   MEDLINE  22835951
   PUBMED   12884007
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..1225
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest of
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector."
ORIGIN

  Query Match             8.0%;  Score 80.2;  DB 9;  Length 1225;
  Best Local Similarity   41.9%;  Pred. No. 7.2e-06;
  Matches  316;  Conservative    0;  Mismatches  439;  Indels    0;  Gaps    0;

Qy        254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
              | || ||||||||    |  |  |||  |  || | | ||||    |||| |
Db       1169 AAAAAAAAAAAAAAAAAAAAANAAAAANAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAA 1110

Qy        314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
               |  || |   ||  ||   ||   | || |||  |  |||   |||  | |   ||  |
Db       1109 AAAAAAAAANNAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1050

Qy        374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                    | || |  | | | |||   |   | ||| |   ||| ||||        |  |
Db       1049 AANNAAAAAAAANAAAANAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAA 990

Qy        434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                  |   ||||      |   | |  | ||  |||| |   |       |||| ||
Db        989 AANNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 930

Qy        494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
               ||||     |   ||| || || ||   ||||||     | || || |  | || ||||
Db        929 NAAAAANAAAANANAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 870

Qy        554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
                 || || | |||||| ||| ||  |   | |||||  ||||   |        |  |
Db        869 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAANAAAAAA 810

Qy        614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
                   |||  |             |   ||  ||     |        ||     ||
Db        809 AAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAANANAAAAAAA 750

Qy        674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
              ||  |      ||| ||  |  ||   |  |    |  |||     ||  ||  |    |
Db        749 AAAAAAAAAAAAAAAAAAAANAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 690

Qy        734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
              ||  |    |   |  ||   | |   |||    |    |    ||     |    |
Db        689 AAAAAAAAAAAAAANNAAANNAAAAAAAAAAAAAANAAAANAAAAAAAAAAAAAAAAAAA 630

Qy        794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
                | ||  | ||| |||  ||     |   ||     |  |||| ||  | ||||| |
Db        629 AAAAAANAAAAAAAAAAAAAAANAAAANANAAAAAAAAAAAAAANAAAAAAAAAAANANN 570

Qy        854 CTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACT 913
                  |   || |||  |||       |||||  ||       | |    | |     |
Db        569 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAANAAAAAANAAAA 510

Qy        914 TATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAAT 973
               |     || |  |        ||       | || |||    |||| ||       ||
Db        509 NAAAANAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 450

Qy        974 TATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
               |   || ||| | |  |||||| | || ||  ||
Db        449 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 415


RESULT 19
CL119201
LOCUS       CL119201                1256 bp    DNA     linear   GSS 05-JAN-2004
DEFINITION  ISB1-76J12_Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-76J12,
            genomic survey sequence.
ACCESSION   CL119201
VERSION     CL119201.1  GI:40612836
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1256)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 75000   Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 233
            High quality sequence stop: 402.
FEATURES             Location/Qualifiers
     source          1..1256
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:8364"
                     /clone="ISB1-76J12"
                   /clone_lib="ISB1"
                   /note="Vector: pBeloBAC11; ISB-1 Xenopus tropicalis BAC
                   Library Segment 1"
ORIGIN

  Query Match              7.9%;  Score 80;  DB 9;  Length 1256;
  Best Local Similarity   37.2%;  Pred. No. 8e-06;
  Matches  281;  Conservative    0;  Mismatches  474;  Indels    0;  Gaps    0;

Qy        254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
              | || ||||||||    |  |  |||  |  || | | ||||    |||| | |
Db        269 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 328

Qy        314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
               |  || |   ||  ||   ||   | || |||  |  |||   |||  | |   ||  |
Db        329 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 388

Qy        374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                    | || || | | | |||   |   | ||| |   ||| ||||        |  |
Db        389 AAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNAAAAAA 448

Qy        434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                 ||   ||||      |   | |  | ||  |||| |   |       |||| ||
Db        449 AAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 508

Qy        494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
               ||||     |   ||| || || ||   ||||||     | || || |  | || ||||
Db        509 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 568

Qy        554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
                 || || | |||||| ||  ||      | ||                     |  |
Db        569 AAAAAAAAAAAAAAAAAAAANAAAAANNNNAAAANNNNNNNNNNNNNNNNAAAAAAAAAA 628

Qy        614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
                   |||  |                                     |
Db        629 AAAAAAAAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNANNNNNNNNNNN 688

Qy        674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
              |   |      ||| ||  |  ||   |  |    |                       |
Db        689 ANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNNNNNNNNNNNNNNNNNNNNNA 748

Qy        734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
              ||  |    |   | |||  ||                       |
Db        749 AAAAAAAAAAAAAAAAAAAAAAANNNNNNNNNNNNNNNNNNNNNNAANANNNNANNNNNA 808

Qy        794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
                | ||  | ||| |||  ||     ||  ||     |  |||   |  | ||||| |
Db        809 AAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAANAAANNAAAAAAAAAAAAAAA 868

Qy        854 CTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACT 913
                  |   || |||  |||       |||||  ||       | |    | |     |
Db        869 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 928

Qy        914 TATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAAT 973
                     || |  |        ||       | || | |    |||| ||       ||
Db        929 ANAAAAAAAAAAAAANAAAAAAAAAAANNAAAAAAAANAAAAAAAAAAAAAAAAAAAAAA 988

Qy        974 TATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
               |   || ||| | |  |||||| | || ||  ||
Db        989 AAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAA 1023


RESULT 20
CL694436/c
LOCUS       CL694436                 776 bp    DNA     linear   GSS 10-JUL-2004
DEFINITION  PRI0164b_A02_2 - PRI0164b.BR (776) Mixed stage fosmid library of P.
            pacificus var. California Pristionchus pacificus genomic, genomic
            survey sequence.
ACCESSION   CL694436
VERSION     CL694436.1  GI:50216344
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 776)
  AUTHORS   Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
  TITLE     AppaDB: an AcedB database for the nematode satellite organism
            Pristionchus pacificus
  JOURNAL   Nucleic Acids Res. 32 (1), D421-D422 (2004)
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            This library was generated at Caltech, Pasadena, USA and end
            sequenced at Vancouver, Canada.
            Seq primer: T7
            Class: fosmid ends.
FEATURES             Location/Qualifiers
     source          1..776
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Mixed stage fosmid library of P. pacificus
                     var. California"
                     /note="Vector: pEpifos-5 Fosmid vector"
ORIGIN

  Query Match              7.9%;  Score 79.6;  DB 9;  Length 776;
  Best Local Similarity   51.4%;  Pred. No. 9.4e-06;
  Matches  184;  Conservative    0;  Mismatches  174;  Indels    0;  Gaps    0;

Qy        199 TCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAAC 258
              || || |  |||||| | |  ||  | || || |  ||  | || || || || ||||||
Db        775 TCGTTCGCCGATGCGCTGGCCGCCGTGGACGTCAGCCCGATCGGTATTGCCGACGATAAC 716

Qy        259 AAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGA 318
               |   ||| || ||  |      | |  | |   |  |   |   |     || || |||
Db        715 GATGCAAAACGCATCCTGCCCGAAGTGCGTGCGCACCTGAAACCGTGGCAGTCCGTCGGA 656

Qy        319 ACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCT 378
              || ||   ||| || | | | ||||  ||       || ||||||||  | || |||||
Db        655 ACGCGCGCGCAGCCGAGCCTGGAAGCCATTGCCGCTCTGAAACCAGACCTGATCATTGCC 596

Qy        379 GATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAA 438
              || |  | | | ||    ||  |||| |  | |||  |  |||| || ||     |
Db        595 GACAGCAGTCGCCATGCGGGGGTTTACATCGCCTTGCAGCAAATCGCGCCGGTACTGCTG 536

Qy        439 CTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAA 498
              || ||      | | | |   ||   ||||||| |  |  ||      |  ||     ||
Db        535 CTTAAGTCCCGCAACGAAACCTACGCTGAAAATTTGCAATCTGCGGCTATCATCGGCGAA 476

Qy        499 GCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
                  | |||||| |   |||     |   |||| | ||| ||||  |  |||  || |
Db        475 ATGGTGGGTAAAAAGCGAGAGATGCAGGCACGCCTGGAACAACATAAAGAGAGGATGG 418


RESULT 21
CG755650/c
LOCUS       CG755650                1231 bp    DNA     linear   GSS 24-OCT-2003
DEFINITION  P051-1-G09.yc Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG755650
VERSION     CG755650.1  GI:37982392
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 1231)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
  TITLE     An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
   MEDLINE  22835951
    PUBMED  12884007
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 004970716013\71
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..1231
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest of
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector."
ORIGIN

  Query Match             7.9%;  Score 79.6;  DB 9;  Length 1231;
  Best Local Similarity   41.1%;  Pred. No. 9.6e-06;
  Matches  310;  Conservative    0;  Mismatches  445;  Indels    0;  Gaps    0;

Qy          254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
                | || ||||||||    |  |  |||  |  || | | ||||    |||| | |
Db         1228 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1169

Qy          314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
                 |  || |   ||  ||   ||   | || |||  |  |||    ||  | |   ||  |
Db         1168 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNAAANAAAAAAAAAAA 1109

Qy          374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                      | || ||   | | |||   |   | ||| |   ||  | ||        |  |
Db         1108 AAAAAAAAAANAAANAAAAAAAANAAAAAAAAAAAAANAAAANAANAAAAAANAAAAAAA 1049

Qy          434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                   ||    |||      |   | |  | ||  |||| |   |       |||| ||
Db         1048 AAAAAAAANAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 989

Qy          494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
                 ||||     |   ||| || || ||   ||||||     | || || |  | || ||||
Db          988 AAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAA 929

Qy          554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
                   || || | |||||| ||| ||  |   | ||||   ||||   |        |  |
Db          928 ANNAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAANAAAAAAAAAAANANAAAAAAAAA 869

Qy          614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
                     |||  |             |   ||  ||     |        ||     ||
Db          868 AAAAAAAAAAAAAAAAANNAAANNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANN 809

Qy          674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
                    |       || ||  |  ||   |  |    |  |||     ||  ||  |    |
Db          808 NNAAANNAAANNAAAAAAAAAAAAAAAAAAAAAAAAANAAANNAAAAAAAAANAAANAAA 749

Qy          734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
                ||  |    |   | |||  || |   |||         ||   ||     |    |
Db          748 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAANAAAAAAAAAAAAAAAAAAAAAN 689

Qy          794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
                  | ||  | ||| |||  ||     ||  ||     |  |||| ||  | |||||
Db          688 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAN 629

Qy          854 CTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACT 913
                    |   || |||  |||       |||||  ||       | |    | |     |
Db          628 NNAAAAAAAAAAAAAAAAAAANNAAAAAAAAAAAAAAAAAANAAAAANNAAAANAAAAAA 569

Qy          914 TATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAAT 973
                 |     || |  |        ||       | || |      |||| ||       ||
Db          568 AAAAAANAAAAAAAANANANAAAAANNAAAAAAAAAANNAAAAAAAAAAAAAAAAAAAAA 509

Qy          974 TATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
                 |   || ||| | |  |||||| | || ||  ||
Db          508 AAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAA 474


RESULT 22
AG332167/c
LOCUS       AG332167                1251 bp    DNA     linear   GSS 02-JUN-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-122E04.T7, genomic survey
            sequence.
ACCESSION   AG332167
VERSION     AG332167.1  GI:47905477
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus
  ORGANISM  Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     BAC end Sequences of Library MSMg01
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1251)
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            Tsukuba Institute, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : T7
            LIBRARY
            Vector     : pBACe3.6
            R.Site 1   : EcoRI
            R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
     source          1..1251
                     /organism="Mus musculus molossinus"
                     /mol_type="genomic DNA"
                     /sub_species="molossinus"
                     /db_xref="taxon:57486"
                     /clone="MSMg01-122E04.T7"
                     /sex="male"
                     /tissue_type="mixture of kidney and spleen"
                     /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN

  Query Match             7.9%;  Score 79.6;  DB 9;  Length 1251;
  Best Local Similarity   51.0%;  Pred. No. 9.6e-06;
  Matches  175;  Conservative    0;  Mismatches  168;  Indels    0;  Gaps    0;

Qy          254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
                || | |||| |||    || | | ||  |   |   |||||| |  ||| || |
Db         1192 ATTATAAAATAAAAAAAATAAATNAAAAAAATAATAATAAAAATNAAAATTAAAAAAAAA 1133

Qy          314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
                ||  |  |  |||  ||  ||| ||| |  ||   |  |||  |||   | | | || ||
Db         1132 TATAATAATATAATAAAAATAAATTANATAAANANAAAAAATATAATATAAAATAAAATA 1073

Qy          374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                |     |  |||| | | | ||    |  |||| | |   ||||||||        |  |
Db         1072 TAAAAAAATATAAAAAAAATAATTAAAAATATATATAAAAAAATAAAAAAATAAAAATAA 1013
```

```
Qy         434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                  ||   || |  ||  |   | |  | ||  |||| |   |       |||| ||
Db        1012 AAAAATATAATAAATTAAAAATAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAN 953

Qy         494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
                ||||     |   ||| || || ||   ||||||     | |  || |  | || ||||
Db         952 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAA 893

Qy         554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAA 596
               |  || || | || ||| ||| ||  |   | |||||  ||||
Db         892 TAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAA 850


RESULT 23
CNS036A2
LOCUS       CNS036A2                 625 bp    DNA     linear   GSS 01-SEP-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
            215D15 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL229763
VERSION     AL229763.1  GI:7888758
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetradontoidea; Tetraodontidae; Tetraodon.
REFERENCE   1
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Estimate of human gene number provided by genome-wide analysis
            using Tetraodon nigroviridis DNA sequence
  JOURNAL   Nat. Genet. 25 (2), 235-238 (2000)
  MEDLINE   20296633
   PUBMED   10835645
REFERENCE   2
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
            Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
            Saurin,W., Bernot,A. and Weissenbach,J.
  TITLE     Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Genome Res. 10 (7), 939-949 (2000)
  MEDLINE   20359837
   PUBMED   10899143
REFERENCE   3  (bases 1 to 625)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1..625
                     /organism="Tetraodon nigroviridis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:99883"
                     /clone="215D15"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0AG215CB08SP1-end :
                     PUC-Ori"
ORIGIN

  Query Match             7.9%;  Score 79.2;  DB 9;  Length 625;
  Best Local Similarity   48.2%;  Pred. No. 1.1e-05;
  Matches  201;  Conservative    6;  Mismatches  210;  Indels    0;  Gaps    0;

Qy         180 TGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGT 239
               | || || || ||   | ||  ||| ||  |   ||  ||  |||  ||  :| |  |
Db         209 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTATAAWATATATTA 268

Qy         240 TGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGG 299
               |   |||    |: | || : |||||||  : || |: ||   |||  : | ||||
Db         269 TAAAATAATAAAWTAAAAAWKAAAAAATDTWTTTTTWTAATTTTTATTWAAAAAAAAAAA 328

Qy         300 AAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAA 359
               |||| | |       |  || |   ||  ||   ||   | || |||  |  |||   ||
Db         329 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 388

Qy         360 ACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAA 419
               |  | |   ||  |      | || || | | | |||   |   | ||| |   ||| ||
Db         389 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 448

Qy         420 AATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGC 479
               ||        |  |   ||   ||||      |   | |  | ||  |||| |   |
Db         449 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 508

Qy         480 TTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGA 539
                   |||| ||    ||||     |   ||| || || ||   ||||||     | || |
Db         509 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 568

Qy         540 ACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAA 596
               | |  | || ||||   || || | |||||| ||| ||  |   | |||||  | ||
Db         569 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAA 625


RESULT 24
CNS017Y0/c
LOCUS       CNS017Y0                1101 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC
            BACN37H18 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL108642
VERSION     AL108642.1  GI:5628946
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACN37H18"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : T7"
ORIGIN

  Query Match             7.9%;  Score 79.2;  DB 9;  Length 1101;
  Best Local Similarity   48.6%;  Pred. No. 1.1e-05;
  Matches  174;  Conservative   22;  Mismatches  160;  Indels    2;  Gaps    1;

Qy         239 TTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTG 298
               ||   |||| |||  | || |||||||||  :| ||   ||  |  || | ||||||
Db        453 TTAAAATAGAGGAAAAAAAAAAAAAAATAAWAATAAATAAAWAAAAAAAATATAAAAAAA 394

Qy        299 GAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTA 358
               |||| | |    :  |  |  |   ||  ||  :||   | || |||  |  ||    |
Db        393 AAAAAAAAAAAAAWAAAAAAAGAATAAAAAAAAAWAAWAAAAAAAAAAAAATAAATAAAA 334

Qy        359 AACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATA 418
              ||    |   ||  |    | | || || | | | |||  ||   : ||| | :|||| |
Db        333 AAWATAAAAAAAAAAAAATTAAAAAAAAAAAAAAAAAAAATAAAAWAAAAAAAWTAAAAA 274

Qy        419 AAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATG 478
              |||        |  |   ||    ||| ::|  |     |::|:||  |||| : : |
Db        273 AAAAAAAAWAAAAAAAATAAAAATAAAAWWTAAAAAATAAWWAWAAAAAAAAAWAWAAAA 214

Qy        479 CTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAG 538
               :  :|||| || |  ||||     |  :||| :|  | ||   |||| |     |  ||
Db        213 AWAAWAAAAAAAAT--AAAATAAAAAAAWAAAAWAATAAAAAAAAAAATAAAAAAAATAG 156

Qy        539 AACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAA 596
              || |  |: | |||| |||  ||::||||||| ||:| |  ||  | |||||: |||:
Db        155 AAAAAAAWTAAAAAAATGATAAAWWTAAAAAAAAAWTTAAAATAAAAAAAAAWAAAAW 98


RESULT 25
CG744812/c
LOCUS       CG744812                1353 bp    DNA     linear   GSS 24-OCT-2003
DEFINITION  P037-3-B03.ya Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG744812
VERSION     CG744812.1  GI:37965680
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 1353)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
  TITLE     An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
   MEDLINE  22835951
   PUBMED   12884007
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..1353
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest of
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector."
ORIGIN

  Query Match             7.9%;  Score 79.2;  DB 9;  Length 1353;
  Best Local Similarity   40.2%;  Pred. No. 1.2e-05;
  Matches  306;  Conservative    0;  Mismatches  455;  Indels    0;  Gaps    0;

Qy        244 ATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAA 303
              | |    |  | |  |||  |||    |  |  |||  |  || | | ||||    ||||
Db       1014 AAAANNNANAAAANNAAANNAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAA 955

Qy        304 TACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCA 363
               | |       |  || |   ||  ||   ||   | ||  ||  |  |||   |||
Db        954 AAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAANNAAAAAAAAAAAAAAAAAAN 895

Qy        364 GATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATT 423
               |   ||  |      | || || | | | |||   |   | | | |   ||| ||||
Db        894 AAAAAAAANAANNAAAAAAAAAAAAAAAAAAAANNAAAAAAAANANAAAAAAAAAAAAAN 835

Qy        424 GCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTT 483
                    |  |        ||||      |   |    | ||  |||| |   |
Db        834 ANAAAAAANAAANNNAAAAAAAAAAAAAAAAAAANAAAAAAANAAAAAAAAAAAAAAAAA 775

Qy        484 AAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACAC 543
              |||  |     ||||     |   ||| || || ||   ||||||     | || || |
Db        774 AAANAANAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 715

Qy        544 GATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTG 603
               | || || |   || || | |||||| ||| ||  |   |  ||||  ||||   |
Db        714 AAAAAAAANAAAAAANAAAAAAAAAAAAAAAAAANAAAAAAANAAAAAAAAAAAAAANNA 655

Qy        604 CCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAA 663
                   |  |      |||  |             |   ||  ||     |        ||
Db        654 AANNAAAAAAAAAANAAAANAAAAAAAAAAAANAAANAAAAAAAAAAAAANAAAAAANAA 595

Qy        664 TTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGT 723
                   ||   |   |      ||| ||  |  ||   |  |    |  |||     ||
Db        594 AAAAAAANANANAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAANNAAAAAAAAAAAA 535

Qy        724 AAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCT 783
              ||  |    |||  |    |   | |||  || |   ||     |    ||   ||
Db        534 AAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAANAAAANAAAAAANNAAAANNAAAAAA 475

Qy        784 GAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTA 843
               |    |     | ||  | ||| |||  ||     ||  ||     |  |||| ||  |
Db        474 AAAAAAAAAAAAAAAAANAAAAAAAAANAAAAANAAAAAAAAAAAAAAAAAAAAAAANAA 415

Qy        844 GAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTA 903
               ||||| |      |   || |||  |||       |||||  ||       | |    |
Db        414 AAAAAAAAAAAANNAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 355

Qy        904 GACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAA 963
                     |   |     || |  |         |       | || |||    |||| ||
Db        354 ANNAAAAAAAAAAAAAAAAAANAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAA 295

Qy        964 CTTGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATAA 1004
                     ||  |   |  ||| | |  |||||| | || ||
Db        294 AAAAAANAANAANNAANAAAANAAAAAAAAAAAAAAAAAAA 254


RESULT 26
CG757503/c
LOCUS       CG757503                1392 bp    DNA     linear   GSS 24-OCT-2003
DEFINITION  P052-4-C08.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG757503
VERSION     CG757503.1  GI:37986131
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 1392)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
  TITLE     An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
   MEDLINE  22835951
   PUBMED   12884007
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..1392
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest of
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector."
ORIGIN

  Query Match             7.8%;  Score 79;  DB 9;  Length 1392;
  Best Local Similarity   44.9%;  Pred. No. 1.3e-05;
  Matches  338;  Conservative    0;  Mismatches  412;  Indels  · 3;  Gaps    1;

Qy         254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
               | || | |||||     || | ||||    ||| |  ||||  |  || |   | || |
Db         928 AAAATATAAAAANAAAAATAAATAAAATTATAAAAATTAAATATAAAATAATAAATTATA 869

Qy         314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
               |   || |   ||  ||   || ||| || | ||   | || | ||   | |    |
Db         868 TTATAAAAAAAAAAAAAAAAAAATTAAAAAATATATATTAATTAAATAAATAAAATAAAT 809

Qy         374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                     |  | || | | | |||   || |||| | |   | ||||          |  |
Db         808 AATAAAAATAAAAAAAAAATAAAAAAATATATATAAATAAATATAATTAATTATTAAATA 749

Qy         434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                  ||   ||||   |  ||  | |  | ||| |||||||  |       |||  || |
Db         748 AATAAAAAAAAAAAATAAATAAAAAAAAAAATAAAAATATAAAAAAAAAAAAATAAAATA 689

Qy         494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
                |||   |  |   ||   | |  ||   |  |||   | | || || |  |  | ||||
Db         688 AAAATATTAAATAAAATATAAATAAATAAATTAAAATATAAAAAAAATAAAAATAAAAAA 629

Qy         554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
               |  || | |   ||||| ||| ||  ||  |||| ||  |||    |          ||
Db         628 TAAAATATTTATAAAAATAAAAAAAAATNTATAATAAAAAAATAAAAAAAAAAAATATAA 569

Qy         614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
                   |     |    |   |  | ||   |   | | |  |  |  | ||||    |
Db         568 AAATTTTTAAATAAATTTATATTAATTTTATTTATTTTATTAATTATTAATTATATATAT 509

Qy         674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
                  ||   | || | ||  |  ||   |  ||  |   |||     |  ||| ||    |
Db         508 TTATATAATATATAAAAAAAAAAAAAAAAAATAATTAAAAAAAAAAATATAATTAAAAAA 449

Qy         734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
               ||  |   ||   | |||  |  | | |||     |   ||    |     |   ||| |
Db         448 AATAATAATAAAAAAAAAAAATAAATAAAATAAATTAAAAATAATAATAAAATATTATAT 389

Qy         794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
               | | ||| | ||| |||  || |   ||  |   |  |  | || ||  | | | | |
Db         388 TTAAAATAAAAAATAAATAAAATAAAAATTATAATATAAAATAAAAAAAAAATATATAAT 329

Qy         854 CTGTATGGA---AGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTG 910
                  | |  |   |  | |  ||   | | ||     ||  | |  |||||  | |
Db         328 TAATTTATATTTATTATTATAAATATATAAATTTATAAATTATATATATTAAAAATTTAT 269

Qy         911 ACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTG 970
               |   ||    || |     |   | |||||      || |||    |||| |  |   |
Db         268 ATAAATATATAAAAAATATTAAATAAATTTAAAATAAAAAAAAAAAAAAAAATATATATT 209

Qy         971 AATTATCTAAGAAAGATAGTAAAAAAGATAATA 1003
               |||| | ||| |||  |   ||||||  ||| |
Db         208 AATTTTTTAATAAAATTTAAAAAAAAATTAAAA 176


RESULT 27
CL078538
```

```
LOCUS       CL078538                1536 bp    DNA     linear   GSS 31-DEC-2003
DEFINITION  CH216-151C11_Sp5.1 CH216 Xenopus tropicalis genomic clone
            CH216-151C11, genomic survey sequence.
ACCESSION   CL078538
VERSION     CL078538.1  GI:40534451
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1536)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: Sp5 atctgccgtttcgatcct
            Class: BAC ends
            High quality sequence start: 1061
            High quality sequence stop: 1145.
FEATURES             Location/Qualifiers
     source          1..1536
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-151C11"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"
ORIGIN

  Query Match             7.8%;  Score 79;  DB 9;  Length 1536;
  Best Local Similarity   51.3%;  Pred. No. 1.3e-05;
  Matches  178;  Conservative    0;  Mismatches  169;  Indels    0;  Gaps    0;

Qy         254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
               | || ||| ||||    |  |  |||  | ||| | | ||||    ||||   |
Db         912 AAAAAAAATAAAAAAAAAAAAAAAAATAAATAAAAAAAAAAAAAAAAAAAANAAAAAAAA 971

Qy         314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
                |  || |   ||  ||   ||   | || |||  |  ||| | |||  | | | ||  |
Db         972 AAAAAAAATAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAATAAAANA 1031

Qy         374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                     | || || | | | |||   |   | ||| |   ||| ||||        |  |
Db        1032 AAAAAAAAAAAAAAAAAAAAAAAATAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1091

Qy         434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                | ||   ||||      |   | |  | ||  |||| |   |       |||| ||
Db        1092 ATAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1151

Qy         494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
                ||||   | |  |||| || || ||   ||||||     | || || |  | || ||||
Db        1152 AAAAAAAATAATATAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1211

Qy         554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTA 600
                  || || |||||||| ||| ||  ||  | |||||| ||||  ||
Db        1212 AAAAAAAAAATAAAAAAAAAAAAAAAAATAAAAAAAAATAAAAAAATA 1258


RESULT 28
CL077122
LOCUS       CL077122                1162 bp    DNA     linear   GSS 31-DEC-2003
DEFINITION  CH216-143E5_Sp6.1 CH216 Xenopus tropicalis genomic clone
            CH216-143E5, genomic survey sequence.
```

```
ACCESSION   CL077122
VERSION     CL077122.1  GI:40533035
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1162)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 757
            High quality sequence stop: 808.
FEATURES             Location/Qualifiers
     source          1..1162
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-143E5"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"
ORIGIN

  Query Match             7.8%;  Score 78.8;  DB 9;  Length 1162;
  Best Local Similarity   45.5%;  Pred. No. 1.4e-05;
  Matches  316;  Conservative    0;  Mismatches  378;  Indels    1;  Gaps    1;

Qy        254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
              | || ||||||||    |  |  |||  |  || | | ||||    |||| | |
Db        469 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAA 528

Qy        314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
               |  || |   ||  ||   ||  || || |||  |  |||   |||  | ||  ||  |
Db        529 AAAAAAAAAAAAAATAAAAAAAAATATAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAA 588

Qy        374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                      || || | | | |||   |  || ||| |   |||||||||       |
Db        589 AAAAAATAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAATAAAATAAAATAAAAAT 648

Qy        434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                 ||   ||||      |   | |  | ||  |||| |   |       |||| || |
Db        649 AAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATA 708

Qy        494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
               ||||     |   ||| || ||  |  |||||||     | || || |  | || ||||
Db        709 AAAAAAAAAAAAAAAAAAAAAAATNATATAAAAAAAAAAAAAAAAAAGAAAAAAAAAAAA 768

Qy        554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
                 || || |||||||| ||| ||  |   | |||||  ||||         |  | ||
Db        769 AAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAATATAAAATAT 828

Qy        614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
                   |||  |             |   ||   |     |        |||    ||
Db        829 AAAAAAAAAAAAAAAAAAAAAAAAAAATAAAATAAAAAAAAAAAAAAAAATAAAAAAAAA 888

Qy        674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
              ||  |    | ||| ||  | ||| | |  ||  ||  | |     |    |  |   ||
Db        889 AAAAAAAAATAAAAAAATAAATAAATAAAAATAATAAAATAAAAAAATAATAAAAAAATA 948

Qy        734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
              ||  |    | | | |||| || |   |||    || | ||   ||     |    |
Db        949 AAAAAAAAAAATAAAAAATAAAAAAAAAAAAAAAATATAAAAAAAAAAAAAAAAATAAAAA 1008

Qy        794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
                | ||  | ||| |||  ||     ||  ||     |  | || || || || | |||
Db       1009 AAAAAAAAATAAAAAAAAAAATAAGAAAAAAATAAAAATAATAAAAAATA-AATATGATA 1067

Qy        854 CTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACT 913
                 ||   ||  ||| | |     | |||| | ||     | |||  |||||   | | |
Db       1068 AAATAAAAAAAGAATAATAATAAATAAAAATTAAAAAAAATAATAAATTAGAAAATAAAT 1127

Qy        914 TATGGGCAAGATCACGTGGTTTAATTTCTTCAGAA 948
               |     || |  |  |     ||| | |  | ||
Db       1128 AAGAATAAAAAAAAAATAAAAAAATATATAAAAAA 1162


RESULT 29
BF273407/c
LOCUS       BF273407                1185 bp    mRNA    linear   EST 07-MAR-2001
DEFINITION  GA__Eb0017O23f Gossypium arboreum 7-10 dpa fiber library Gossypium
            arboreum cDNA clone GA__Eb0017O23f, mRNA sequence.
ACCESSION   BF273407
VERSION     BF273407.1  GI:11204477
KEYWORDS    EST.
SOURCE      Gossypium arboreum
  ORGANISM  Gossypium arboreum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE   1  (bases 1 to 1185)
  AUTHORS   Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J.,
            Henry,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
  TITLE     An integrated analysis of the genetics, development, and evolution
            of the cotton fiber
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: TAATACGACTCACTATAGGG
            High quality sequence start: 540
            High quality sequence stop: 1175.
FEATURES             Location/Qualifiers
     source          1..1185
                     /organism="Gossypium arboreum"
                     /mol_type="mRNA"
                     /strain="AKA"
                     /cultivar="8400"
                     /db_xref="taxon:29729"
                     /clone="GA__Eb0017O23f"
                     /tissue_type="Fibers isolated from bolls harvested 7-10
                     dpa"
                     /lab_host="E. coli"
                     /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
                     /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
ORIGIN

  Query Match             7.8%;  Score 78.8;  DB 2;  Length 1185;
  Best Local Similarity   39.5%;  Pred. No. 1.4e-05;
  Matches  251;  Conservative    0;  Mismatches  384;  Indels    0;  Gaps    0;

Qy        254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
              | || ||||||||    |  |  |||  |  || | | ||||    |||| | |
Db       1164 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1105

Qy        314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
               |  || |   ||  ||   ||   | || |||  |  |||   |||  | |    |  |
Db       1104 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAANAAAA 1045

Qy        374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
```

```
                     | || || | | | |||   |   | ||| |   ||| ||||        |  |
Db       1044 AAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAA 985

Qy        434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                 ||   ||||      |   | |  | ||  |||| |   |       |||| ||
Db        984 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 925

Qy        494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
               ||||     |   ||| || || ||   ||||||     |  | || |  | || ||||
Db        924 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNAAAAAAAAAAAAAAAAAAAA 865

Qy        554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
                 || || | |||||| ||| |   |   | |||||  ||||   |        |  |
Db        864 AAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAA 805

Qy        614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
                                      |   |                   |
Db        804 AAAANNNNNNNNAANNNNANANNAAAANANNNNNNNNANNNNNNNNNAANANNNANNNNN 745

Qy        674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
                         ||| ||  |  ||   |          ||      ||  ||       |
Db        744 NNNANNNAAAAAAAANAAANANNAANNNANNNNANNNNAAANAAAAAAAAAAAAAANNAAAAA 685

Qy        734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
              |   |    |   | |    || |   |||    |    ||   ||     |    |
Db        684 ANNAAANNAAANNAAANNAAAAAAAAAAAAAAAANNAAAAAAAAAAAAAANNAAAAANAAAAA 625

Qy        794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
                | ||  | ||| |||  ||     ||  ||     |  |||| ||  | ||||| |
Db        624 AAAAAAAAAAAAAAAAAAAANAAAANAAAAAAAAAAAAAAAANNAAAAAAAAAAAAAAAAAAAAN 565

Qy        854 CTGTATGGAAGAAATTAAACGCTGTGAAAAATCAA 888
                  |   || |||  |||       | |||  ||
Db        564 AAANAAAAAAAAAAAAAAAAANAAAAAAAAANAAANNAA 530

RESULT 30
CL078190/c
LOCUS       CL078190                1217 bp    DNA     linear   GSS 31-DEC-2003
DEFINITION  CH216-148G2_Sp6.1 CH216 Xenopus tropicalis genomic clone
            CH216-148G2, genomic survey sequence.
ACCESSION   CL078190
VERSION     CL078190.1  GI:40534103
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1217)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 229
            High quality sequence stop: 312.
FEATURES             Location/Qualifiers
     source          1..1217
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-148G2"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"
ORIGIN

  Query Match             7.8%;  Score 78.8;  DB 9;  Length 1217;
  Best Local Similarity   45.0%;  Pred. No. 1.4e-05;
  Matches  342;  Conservative    0;  Mismatches  412;  Indels    6;  Gaps    1;

Qy        251 ATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTT 310
              || |||| ||||||||    |  |  |||  |   |   | ||||    |||| | |
Db       1077 ATAATAAAAAAAAAAAAAAAAAAAAAAAAATAAAATAATAAAAAAAAAAAAAAAAAAAAAAA 1018

Qy        311 CTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAA 370
               | ||  || |   ||  ||   ||   | |   ||  |  |||  ||||  | |  |||
Db       1017 ATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATATAAAAAAAAAAAAATAAAAAAAAAAATAA 958

Qy        371 TTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTA 430
                |      | || || | | | |||   |   | ||  | | ||| ||||        |
Db        957 AAAAATAAAAAAAAAAAAAAAAATAAAAAAAAAAGAAAATAAATAAAAAAAAAAAAAAAAAAA 898

Qy        431 CGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAA 490
                ||  ||   ||||   |  |   | || | ||  |||| |        |  |||| ||
Db        897 AAATAAAAAAAAAAAAAAAATAAACATAAATAAAAAAAAAAAAAAAAAAGAAAATAAAAAAAAA 838

Qy        491 TTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGA 550
                  ||||     |   ||| || || ||   ||||||     | ||||| |  | || |
Db        837 AAAAAAAAAAAAACAAAAAAAAAAAAAACAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAA 778

Qy        551 AAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAG 610
              |||   ||||| | |||||| ||| ||  |   | |||||  ||||   |        |
Db        777 AAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAAACAAAAAAAAAGAA 718

Qy        611 TAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAA 670
               |      |||  |             | | ||  |||    |        ||     ||
Db        717 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAAAAAAAACAACAAAAAAAAAAAAAAAAAAAA 658

Qy        671 GTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATC 730
                 ||  |      ||| ||  |  ||   |  |    |  |||     ||  ||  |
Db        657 AAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 598

Qy        731 TTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTA 790
                 ||  |    |   | |||  || |   |||    |    ||   ||     |
Db        597 AGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCACGACC 538

Qy        791 TGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTC------ACTAAAAGAACTAG 844
               |  || ||  | ||| |||  ||     ||  |||          |  |||| ||  |
Db        537 CGCCCAAAAAAAAAAAAAAAAAAAAAGGAAAAAAAACAAAGGAGAAAAAAAAAAAAAAAAAA 478

Qy        845 AAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAG 904
              |||||||      | | |  |||  ||| |  |   ||||  ||      | | ||||
Db        477 AAAAAGAAAAAAAAAGGAACAAAACAAAAAGGGGGAGAAAAAAAAAAAAAGGGGTTTTTTTT 418

Qy        905 ACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAAC 964
                  |     |     || |    | |||||  | | |||  || |||    ||||
Db        417 TTTTTCCAAAAAAAAAAAAAAGGGGGGGGTTTTTTCTTTTCTTAAAAAAAAAAAAAAAATTGT 358

Qy        965 TTGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATAA 1004
              ||   |||  |   ||  || | |  | |||| | || ||
Db        357 TTTACGAACAAATGAAATAAAAAAAAAATAAAAAAAAAAAA 318

RESULT 31
CG754010/c
LOCUS       CG754010                1276 bp    DNA     linear   GSS 24-OCT-2003
DEFINITION  P049-2-C03.ya Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG754010
VERSION     CG754010.1  GI:37979071
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
```

```
REFERENCE   1  (bases 1 to 1276)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
  TITLE     An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
  MEDLINE   22835951
   PUBMED   12884007
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..1276
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest of
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector."
ORIGIN

  Query Match             7.8%;  Score 78.8;  DB 9;  Length 1276;
  Best Local Similarity   40.4%;  Pred. No. 1.4e-05;
  Matches  305;  Conservative    0;  Mismatches  450;  Indels    0;  Gaps    0;

Qy        254 ATAACAAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
              | || ||||||||    |  |  |||  |  || | | ||||    |||| | |
Db       1048 AAAAAAAAAAAAAANAAAAAAANAAAANAANAAAAAAAAAAAAAAAAAAAAAAAAANAAA 989

Qy        314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
               |  || |   ||  ||   ||   | || |||  |  | |   |||  | |   ||  |
Db        988 AAAAAAAAAANAAAAAAANAAAAAAAAAAAAAAAAAAAANAAANAAAAAANAAAAAAAAA 929

Qy        374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                    | || |  | | | |||   |   | |||     ||  ||||        |  |
Db        928 AAAANNAAAAAANAAAAAAAAAAANAAAAAAAAAAANAAAAANAAAAAAAAAAAAAAAAA 869

Qy        434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                 ||   ||||      |   | |  | ||  | ||     |       |||| ||
Db        868 AAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAANAAANAAAAAANAAAAAAAAAAAAAA 809

Qy        494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
               ||||     |   ||| || || ||   ||||||       || || |  | || ||||
Db        808 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAA 749

Qy        554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
                 || || | |||||| ||| ||  |   | |||||  ||||   |        |  |
Db        748 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAANAAAAAAAA 689

Qy        614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
                   |||  |             |   ||  ||     |        ||     ||
Db        688 AAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAANAAAAA 629

Qy        674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
              ||  |      ||| |   |  ||   |  |    |    |     ||  ||  |    |
Db        628 AAAAAAANAAAAAANANAAAAAAAAAAAAAAANNAAAANNAAAANNAAAAAAAAAAAAAA 569

Qy        734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
              ||  |    |   | |||  || |   |||    |    ||   ||     |    |
Db        568 AAAAAAAANAAAAAAAAAAAAAAAAANAAAAAAAAANNAAAANNAAAAAAAAAAAAAAAA 509

Qy        794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
                | ||  | ||  |||  ||     ||  ||     |  ||||     | ||||| |
Db        508 NAAAAAANAAAANNAAAAAAAAAAAAAAAAAAAAAAAANNAAAAANNAAANAAAAANAAA 449

Qy        854 CTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACT 913
                  |   || ||   ||        |||||  ||       | |    | |     |
Db        448 AAAAAAAAAANAANAAAANNNNANAAAAAAAAAAAAAAAAAAAAAAAAAAAAANANAAAA 389

Qy        914 TATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAAT 973
               |     || |  |        ||       |  | |||    |||| ||       ||
Db        388 AAAAAAAAAAAAAAAANAAAAAAANAAAAANANNANAAANNNAAAAAAAANAAAAAAAAA 329

Qy        974 TATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
               |   ||  |  | |  |||||| | || ||  ||
Db        328 AAAAAAANNANAAAAAAAAAAAAAAAAAAAAAAAA 294


RESULT 32
AG350118/c
LOCUS       AG350118                1472 bp    DNA     linear   GSS 02-JUN-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-146K03.TJ, genomic survey
            sequence.
ACCESSION   AG350118
VERSION     AG350118.1  GI:47923428
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus
  ORGANISM  Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     BAC end Sequences of Library MSMg01
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1472)
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            Tsukuba Institute, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : TJ
            LIBRARY
            Vector     : pBACe3.6
            R.Site 1   : EcoRI
            R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
     source          1..1472
                     /organism="Mus musculus molossinus"
                     /mol_type="genomic DNA"
                     /sub_species="molossinus"
                     /db_xref="taxon:57486"
                     /clone="MSMg01-146K03.TJ"
                     /sex="male"
                     /tissue_type="mixture of kidney and spleen"
                     /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN

  Query Match             7.8%;  Score 78.8;  DB 9;  Length 1472;
  Best Local Similarity   39.6%;  Pred. No. 1.4e-05;
  Matches  299;  Conservative    0;  Mismatches  456;  Indels    0;  Gaps    0;

Qy        254 ATAACAAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
              | || ||||||||    |  |  |||  |  || | | | ||    ||||   |
Db       1102 AAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAANAANAAAAA 1043

Qy        314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
               |  || |   ||  ||   ||   | || |||  |  |||   |||  | |   ||  |
Db       1042 AAAAAAAAAANAAAAAAAAAAANNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 983
```

```
Qy      374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                  | || ||   | | |||   |   | ||  |   ||| ||||        |  |
Db      982 AANANNAAAAAAAANNAAANAAAAAAAAAAAAAANNAAAAAAAAAAAAANANAAAAAAAA 923

Qy      434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
               ||   ||||      |   | |  |  |  |  | |   |       |||| ||
Db      922 AAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAANNAAAAAAAAAAAAAAAAAAAAAAAA 863

Qy      494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
             ||||     |   ||| || || ||   ||||||     | || || |  | || ||||
Db      862 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAANAAAAAAAAA 803

Qy      554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
               || || | ||| || ||| ||  |   | |||||  ||||   |        |
Db      802 AAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAANANNAAANA 743

Qy      614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
                 |||  |             |   ||  ||     |        ||     ||
Db      742 AAAAAAAAAAAAAAAAAAAAAAAAAANNAAAAAAAAAAAAAAAAAAAAAANNAAAAAANN 683

Qy      674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
            ||  |      ||| ||  |  ||   |  |    |  |||     ||  ||       |
Db      682 AANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAANAAAAA 623

Qy      734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
            ||  |    |   | |||   | |   |||    |    ||   ||     |    |
Db      622 AAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 563

Qy      794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
              | ||  | ||| |||  ||     ||  ||     |  |||| |   | ||||| |
Db      562 AAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAA 503

Qy      854 CTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACT 913
                |   || |||  |||       |||||  ||
Db      502 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAANNNNANNNGGNNNNNNNNNNNNNA 443

Qy      914 TATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAAT 973
                                           | || |||    |||| ||       ||
Db      442 ANNNNNNNNNNNNNNNNNNNNNNNNAAAAAAAAAAAAAAAANAAAAAAAANAAAAAAAAA 383

Qy      974 TATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
             |   || ||| | |  |||||| | || ||  ||
Db      382 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 348


RESULT 33
CC268594/c
LOCUS       CC268594                1352 bp    DNA     linear   GSS 13-MAY-2003
DEFINITION  CH261-67F12_Sp6.1 CH261 Gallus gallus genomic clone CH261-67F12,
            genomic survey sequence.
ACCESSION   CC268594
VERSION     CC268594.1  GI:30619779
KEYWORDS    GSS.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 1352)
  AUTHORS   Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
            Warren,W., Graves,T., Mardis,E. and Wilson,R.
  TITLE     Gallus gallus BAC End Reads
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K. Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 182000    Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 19
            High quality sequence stop: 132.
FEATURES             Location/Qualifiers
     source          1..1352
                     /organism="Gallus gallus"
                     /mol_type="genomic DNA"
                     /strain="Red Jungle Fowl"
                     /db_xref="taxon:9031"
                     /clone="CH261-67F12"
                     /sex="female"
                     /cell_line="UCD001, inbred 256"
                     /clone_lib="CH261"
                     /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                     CH261 Female Chicken library - for library and clone
                     ordering information: http://www.chori.org/bacpac"
ORIGIN

  Query Match              7.8%;  Score 78.6;  DB 8;  Length 1352;
  Best Local Similarity   40.5%;  Pred. No. 1.5e-05;
  Matches  306;  Conservative    0;  Mismatches  449;  Indels    0;  Gaps    0;

Qy      254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
            | || ||||||||    |  |  |||  |  || | | ||||    |||| | |
Db     1238 AAAAAAAAAAAAAAAANAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAANAANNA 1179

Qy      314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
             |  || |   ||  ||   ||   |  | |||  |  |||   |||  | |    |  |
Db     1178 AAAAAAAANAAAANAAAAAAAAAAAAANAAAAAAAANAAAAAAAAAAAAAAAAAANANAA 1119

Qy      374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                  | || || | | | |||   |   | ||| |   ||| ||||        |  |
Db     1118 AAAAAAAAAAAAAAAAAAANAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAANANAAAA 1059

Qy      434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
               ||   ||||      |   | |  | ||  ||   |   |       |||| |
Db     1058 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNNAAAAAAAAAANAAAAAAANAAA 999

Qy      494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
             || |     |   ||| || || ||   ||||||     | || || |  | || ||||
Db      998 AAANANAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAANAAAAAAAAAAAAANAAAA 939

Qy      554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
               || || | ||||||  |||     |   | |||||   |||   |        |  |
Db      938 AAAAAAAAAAAAAAAAAAAAAANNANAAANANAAAAAAANAAANAAAAAAAANAAAAAAA 879

Qy      614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
                 |||  |             |   ||  ||     |        ||     ||
Db      878 AAAAAAAAAAAAAAAAAAAANAAAAANAAAAAAANNAAAAAAANAAAAAAAAAAAAANAA 819

Qy      674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
            ||         |||  |  |  |    |  |    |  |||     ||  ||  |    |
Db      818 AAAANAANNAAAAAANANAAAAANNAAAAAAAAAAAAAAAAAAAAAAANAAAAAANAAAA 759

Qy      734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
            ||  |    |   | |||  || |   |||    |    ||    |     |    |
Db      758 AAAAANNNAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAA 699

Qy      794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
              | ||  | |||  ||  ||     ||   |     |  |||| ||  |  |||| |
Db      698 AAAAAAAAAAAAAANAAAAAAAANNAAAANNAAAAAAAAAAAAAAAAAAANNAAAAAAAA 639

Qy      854 CTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACT 913
                |   || |||  |||       |||||  ||       | |    | |     |
Db      638 AAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAANANANAAANAAAAAAAAAAAAAA 579

Qy      914 TATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAAT 973
             |     || |  |        ||       | || |||    |||| ||       ||
Db      578 AAAAAAAAAAAAAANAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 519

Qy      974 TATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
             |   || ||| |    |||||| | || ||  ||
Db      518 AAAAAAAAAAAAANNANAAAAAAAAAAAAAAAAAA 484


RESULT 34
CG753732/c
```

```
LOCUS       CG753732                1811 bp    DNA     linear   GSS 24-OCT-2003
DEFINITION  P048-4-G03.ya Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG753732
VERSION     CG753732.1  GI:37978509
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 1811)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
  TITLE     An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
  MEDLINE   22835951
   PUBMED   12884007
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..1811
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest of
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector."
ORIGIN

  Query Match             7.8%;  Score 78.6;  DB 9;  Length 1811;
  Best Local Similarity   37.4%;  Pred. No. 1.5e-05;
  Matches  282;  Conservative    0;  Mismatches  473;  Indels    0;  Gaps    0;

Qy        254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
              | || ||||||||    |  |   ||  |  || | | ||      |||| | |
Db       1253 AAAAAAAAAAAAAAANNAAAANANAAAAAAAAAAAAAAAANNAAAAAAAAAAAAAAAAAA 1194

Qy        314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
               |  || |   ||   |   ||   |    |||  |  | |   | |  | |   ||  |
Db       1193 AAAAAAAAAAAAAANNAAAAAAAAAAANNAAAAAAAAAANANAAANANAAAAAAAAAAAA 1134

Qy        374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                    | || || | | | |||   |   |   | |   | | ||          |  |
Db       1133 NANNANAAAAAAAAAAAAAAAAAAAAAAAAAANNAAAANNANAAAANNAANANNAAANNA 1074

Qy        434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                 ||   ||||      |   | |  | ||  |||| |   |       ||||  |
Db       1073 AAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNAAAA 1014

Qy        494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
                 ||         ||| |  |  ||   ||||||     | || ||    | || ||||
Db       1013 NNNAAAANNANNAAAAANANAANAAANANAAAAAANANNAAAAAAAANNAAANAAAAAAA 954

Qy        554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
                 || || | || |||  |  ||  |   | |||||  ||||   |        |  |
Db        953 AAAAAAAAAAAAANAAAANANNAAAAAAAAAAAAAAAANAAAAAAAAAAAANNAAAAAAA 894

Qy        614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
                   |||                |    |  ||     |        ||     |
Db        893 AAAAAAAANANNAAANAAANNAAAAAANNAANAAAAAAAAAAAAAAAAAAAAAAAANNAA 834

Qy        674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
              ||  |      ||| ||  |  ||      |    |  |||     |   ||       |
Db        833 AAAAAANNAAAAAAAAAAAAAAAAAAANNAANNAAAAAAAAANNNNANANAAAANNNNNA 774

Qy        734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
              ||  |    |   | |||  || |   ||     |    |    |      |    |
Db        773 AAAAAAAAAAAAAAAAAANNAAAAAAAAANNAAAANNAAANAAAANNAAAAAAAAAAAAA 714

Qy        794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
                | ||  | | | |||   |     ||  ||     |   ||| ||  | ||||| |
Db        713 AAAAAAAAAAANAAAAAAANANANAAAAAAAAAAAAAAAANAAANAAANAAAAAAAAANA 654

Qy        854 CTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACT 913
                      || |||  |||       |  ||   |        | |   | |     |
Db        653 AAAANNNAAAAAAAAAAAAAAAAAAAANNAAAANAAAAAAAAAAAAANAAAAAAANNANA 594

Qy        914 TATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAAT 973
               |     || |  |        ||       |  | | |    |||| ||       ||
Db        593 AAAAANAAAAAAAAAAAAANAAAAAAAAAAAANNAAANAANANAAAAAAANAAAAAAAAN 534

Qy        974 TATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
               |   || ||| | |  |||||| |  | ||  ||
Db        533 AAAAAAAAAAAAAAAAAAAAAAAAANNANAANNAA 499


RESULT 35
CL081338
LOCUS       CL081338                1163 bp    DNA     linear   GSS 31-DEC-2003
DEFINITION  CH216-161H3_RM4.1 CH216 Xenopus tropicalis genomic clone
            CH216-161H3, genomic survey sequence.
ACCESSION   CL081338
VERSION     CL081338.1  GI:40537251
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1163)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: RM4 ctcaagggcatcggtcgagc
            Class: BAC ends
            High quality sequence start: 826
            High quality sequence stop: 926.
FEATURES             Location/Qualifiers
     source          1..1163
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-161H3"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"
ORIGIN

  Query Match             7.8%;  Score 78.4;  DB 9;  Length 1163;
  Best Local Similarity   44.0%;  Pred. No. 1.7e-05;
  Matches  331;  Conservative    0;  Mismatches  422;  Indels    0;  Gaps    0;

Qy        256 AACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTA 315
              || ||||| ||    |     |||  |  || | | ||||    |||| | |       |
Db        258 AAAAAAAANAAAAAAAAAGAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 317

Qy        316 GGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATT 375
               ||| |   ||  ||   ||   | || |||  |  |||   | |  | |   ||  |
Db        318 AGAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAAAAAAGAAAA 377

Qy        376 GCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATT 435
                  | || || | | | |||   |   | ||| |   ||| |||  |  |  |   ||
Db        378 AAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATATTTTTTTTTAATA 437

Qy        436 GAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCA 495
               || | ||||       |    |  | ||  |||| |   |   |||| ||| |     |
Db        438 TAAATAAAAAAAAAGATTTATAAAAAAAAAAAAAAAAAAAAAATTTTTTAAAAACAAAAA 497

Qy        496 AAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATT 555
              |||     |   ||| || || ||   ||||||     | || || |  | |  ||||
Db        498 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAGAAAAAAA 557

Qy        556 GAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCT 615
               || || | |||||| ||| ||  |   | |||||  ||||   |        |  |
Db        558 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 617

Qy        616 GCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAA 675
                 |||  |   |         |   ||  ||   | |        |      ||   ||
Db        618 AAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAGATAAAAAAAAAAATAAAAAAAAAAAA 677

Qy        676 CTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAA 735
                |      ||| ||  |  || | |       |  |||     ||  ||  |    |||
Db        678 AAAAAAAAAAAAAAAAAAAAAAATAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 737

Qy        736 GGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTC 795
                |    |   | ||||||| |   |||    |    ||   ||   | |    |
Db        738 AAAAAAAAAAAAAAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAATAAAAAAAA 797

Qy        796 ATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCT 855
              | ||| | | | |||  ||    |||  ||     |  |||| ||  | ||||| |
Db        798 AAAATAAAAGAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 857

Qy        856 GTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTA 915
                ||  || |||  |||       |||||  ||    |  | |    | |     |   |
Db        858 AAATAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAATTAAAAAAAAAAAAAAAAAAAAAA 917

Qy        916 TGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTA 975
                   || |  |        ||       | || ||||   |||| ||     | ||  |
Db        918 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAATAAAAAA 977

Qy        976 TCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
              |  || |||   |   ||||| | || ||  ||
Db        978 TAAAATAAAATAAAATAAAAAAAAAAAAAAAAA 1010


RESULT 36
CG758143/c
LOCUS       CG758143                1297 bp    DNA     linear   GSS 24-OCT-2003
DEFINITION  P053-3-B11.zb Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG758143
VERSION     CG758143.1  GI:37987385
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 1297)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
  TITLE     An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
  MEDLINE   22835951
   PUBMED   12884007
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..1297
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest of
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector."
ORIGIN

  Query Match             7.8%;  Score 78.4;  DB 9;  Length 1297;
  Best Local Similarity   40.7%;  Pred. No. 1.7e-05;
  Matches  307;  Conservative    0;  Mismatches  448;  Indels    0;  Gaps    0;

Qy        254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
              | || |||| ||     |  |  ||   |  || | | ||||    |||| | |
Db        995 AAAAAAAAANAANAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAANAAAANAA 936

Qy        314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
               |  || |   ||  ||   ||   | || |||  |  |||   |||  | |   ||  |
Db        935 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAANAAAAAAAAAAA 876

Qy        374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                    | || || | |   |||   |   | ||| |   ||| ||||        |  |
Db        875 AAANAAAAAAAAAAAAAANAAAAAANANAAAAAAAAAANAAAAAAAAAAAAAAAAAANAA 816

Qy        434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                 ||   ||||      |   | |  | ||  |||| |   |       |||| ||
Db        815 AAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAANNA 756

Qy        494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
               ||||     |   ||| || || ||   |||| |     | || || |  | || ||
Db        755 AAAAAAAAAAAAAAAAAAAAAAAAAANNAAAAANANAAAAAAAAAAAAAAAAAAAAAANN 696

Qy        554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
                 || || | |||||| ||| ||  |   | |||||  ||||   |        |  |
Db        695 AAAAAAAAAAAAAAAAAAAAAAAANNAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAA 636

Qy        614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
                   |||  |             |   ||  ||     |        ||     ||
Db        635 AAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNAAAAAAA 576

Qy        674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
              ||         ||| ||  |  ||   |  |    |  |||     ||  ||  |    |
Db        575 AAAANAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAANAAAAANNAAAAAAAAAA 516

Qy        734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
              ||       |   | ||   || |   |||    |    ||   ||          |
Db        515 AAAANAAAAAAAAAAAANAAAAAAAAAAAANNAAANAAAAAAAAAAAAAANNAAAAAAAA 456

Qy        794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
                | ||  | ||| |||  ||      |  ||     |  |||| ||  | ||||| |
Db        455 AAAAAAAAAAAAAAAAAAAAANNANANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 396

Qy        854 CTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACT 913
                  |   || ||     |       |||||  ||       | |    | |     |
Db        395 AAANANAAAANAANNNNNAAAAAAAAAAAAAAAAAAAAAAAAAAANNAAAAAAAAAAAAA 336

Qy        914 TATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAAT 973
               |     || |  |        ||       | || |||    | || ||       ||
Db        335 AAAANAAAAAAAAAAAAAANAAAAAAANAAAAAAAAAAAAAANANAAAAAAANNAAAAAA 276

Qy        974 TATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
               |   || ||| | |  |||||| | || ||  ||
Db        275 AAAAAAAAAAAAAAAANAAAAAANANAAAAANNAA 241


RESULT 37
CG744815/c
```

```
LOCUS       CG744815                1380 bp    DNA     linear   GSS 24-OCT-2003
DEFINITION  P037-3-B04.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG744815
VERSION     CG744815.1  GI:37965683
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 1380)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
  TITLE     An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
  MEDLINE   22835951
   PUBMED   12884007
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..1380
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest of
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector."
ORIGIN

  Query Match             7.8%;  Score 78.4;  DB 9;  Length 1380;
  Best Local Similarity   40.4%;  Pred. No. 1.7e-05;
  Matches  256;  Conservative    0;  Mismatches  378;  Indels    0;  Gaps    0;

Qy        254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
              | || ||||| ||    |  |  |||  |  || | | ||||    |||| | |
Db       1210 AAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAA 1151

Qy        314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
               |  || |   ||  ||   ||   | || |||  |  |||   ||   | |   ||  |
Db       1150 AAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAANAAAAAAAANNAAAANNAAANNA 1091

Qy        374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                    | || || | | | |||   |   | ||| |   ||  ||||        |  |
Db       1090 ANANNAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAANNAAAAAAAAAAAAAAAA 1031

Qy        434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                 ||   | |       |   | |  | ||  |||| |   |       || | ||
Db       1030 AAAAANAAANANAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAANAAAANAA 971

Qy        494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
               ||||     |   ||| || || ||   || ||        ||  | |  | || ||||
Db        970 AAAAAAAAAAANAAAAAAAAAAAAAANNAAANAANAAAANNNAANNAAAAAAAAAAAAAA 911

Qy        554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
                 || || | |||||| ||| ||  |   |  ||||  ||||   |        |  |
Db        910 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNAAAAANAAAAAAAAAAAAANNAAAAAN 851

Qy        614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
                   |||  |             |   ||  ||     |        |       |
Db        850 AAAAAAAANNANNAAANNNANNNNANAAAAAAAAAAAAAAAAAAAAAAANAAAAANANAA 791

Qy        674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
              ||  |      ||| ||| |   |   |  |    |  |||     ||  |   |    |
Db        790 AAAAAAAAAAAAAAAAAGGAAANANAAAAAAANNAAAAAAAAAAAAAAAAANNNANANNA 731

Qy        734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
              ||       |   | |||  || |    ||    |    ||   ||     |    |
Db        730 AAAANAANNAAAAANAAAAAAAAAAANNAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAN 671

Qy        794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
                | ||  | ||| |||   |     ||  ||     |  |||  |   | ||||| |
Db        670 AAAAAAAAAAAAAAAAAANNAAAAAAAAAAAAAAAAAAAAAAANAANNNAAAAAAAAANN 611

Qy        854 CTGTATGGAAGAAATTAAACGCTGTGAAAAATCA 887
                  |   || |||  |||       |||||  |
Db        610 AAAAAAANAAAAAAAAAAANAAAAAAAAAAANAA 577


RESULT 38
CL110653/c
LOCUS       CL110653                1594 bp    DNA     linear   GSS 05-JAN-2004
DEFINITION  ISB1-53P23_Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-53P23,
            genomic survey sequence.
ACCESSION   CL110653
VERSION     CL110653.1  GI:40604288
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1594)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 75000   Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 390
            High quality sequence stop: 470.
FEATURES             Location/Qualifiers
     source          1..1594
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:8364"
                     /clone="ISB1-53P23"
                     /clone_lib="ISB1"
                     /note="Vector: pBeloBAC11; ISB-1 Xenopus tropicalis BAC
                     Library Segment 1"
ORIGIN

  Query Match             7.8%;  Score 78.4;  DB 9;  Length 1594;
  Best Local Similarity   46.2%;  Pred. No. 1.7e-05;
  Matches  295;  Conservative    0;  Mismatches  341;  Indels    2;  Gaps    1;

Qy        251 ATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTT 310
              || |||| || ||||||   |  | ||||  |  || | | ||||    |||| | |  |
Db       1432 ATAATAAAAACAAAAATAAAAAAAATAAATAAAAAAAAAAAAAAAAAATAAAAAAAAAAT 1373

Qy        311 CTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAA 370
                     ||     ||   |   || | |  | |||     |||   |||  | |    |
Db       1372 AAAATAAAAATAAAAAAATATAAAAATAATTAAAAAAACAAAAAAAAAAAAAAAAAAATA 1313

Qy        371 TTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTA 430
                | |    | || || | | | |||   | | ||||| | | ||||||||        |
Db       1312 AAAATATAAAAAAAAAAAAAAAAAAAAAAAATAATAAAAAATAAAATAAAAAATAATTAA 1253

Qy        431 CGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAA 490
                |   ||   | || | |  |   | |  ||||  |||| |   |      ||||  ||
Db       1252 AAAAAAAAAAAATAAATATAAAAAAAAA--ATAAATAAAAAAAAAAAAAAAATAAATAAA 1195

Qy        491 TTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGA 550
```

```
              | | ||||   |||   ||| || || ||   || |||     | |  || |  | || |
Db       1194 TATAAAAAAAATTATAAAAAAAAAAAAAAAACAATAAAAAAAAACATAAATAAAAAAAAA 1135

Qy        551 AAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAG 610
              ||||  || || | |||||| ||| ||  ||    | ||| ||||||  |        |
Db       1134 AAATATAAAAAAAAAAAAAAAAAAAAAAAATAATAATAAAACAAAAGAAAAAAAACAAAA 1075

Qy        611 TAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAA 670
               |      |||  |       |   | |   ||  ||     |        ||     ||
Db       1074 AAAAAAAAAAAAAAAAAAAAATAAATAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAA 1015

Qy        671 GTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATC 730
                 ||  |      ||| ||  |  |  | |  |       |||   ||||  ||
Db       1014 AAAAAAAAATAAAAAAATAAAAAAAAGATAAAAAAAAATAAAAAAAATTAAAAAAAACAA 955

Qy        731 TTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTA 790
                |||  |  | | | | |||  |  | | |||    |    ||   ||   |      |
Db        954 AAAAAAAAAATAAATAAAAAAAAATAAATAAAAAAAAAAAAAAAAAAAAAAATACAAACA 895

Qy        791 TGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAG 850
                |  ||||  | ||| |||  ||    |||  ||     |  |||| || || | |||
Db        894 AATAAATAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAATAAATAAAA 835

Qy        851 ATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAA 888
              |      |   || |||  |||   |   |||||  ||
Db        834 AAAAAAAAATAAAAAAAAAAAAAAATAAAAAAAAAAAA 797

RESULT 39
CD388253
LOCUS       CD388253                 971 bp    mRNA    linear   EST 30-MAY-2003
DEFINITION  AGENCOURT_14304648 NIH_MGC_173 Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CD388253
VERSION     CD388253.1  GI:31226341
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 971)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jamie Thompson, University of WI
             cDNA Library Preparation: Gina Zastrow-Hayes
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDKM80  row: j  column: 22
            High quality sequence start: 16
            High quality sequence stop: 510.
FEATURES             Location/Qualifiers
     source          1..971
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic trophoblasts, made from WA01 stem
                     cells"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_173"
                     /note="Vector: pDONR201; Site_1: attP2; Site_2: attP1;
                     LIBR_PRIMING - oligo dT; METHOD - full-length enriched;
                     LIBR_PROVIDER - Bradfield"
ORIGIN

  Query Match             7.8%;  Score 78.2;  DB 6;  Length 971;
```

```
  Best Local Similarity   42.1%;  Pred. No. 1.8e-05;
  Matches  179;  Conservative    0;  Mismatches  246;  Indels    0;  Gaps    0;

Qy        172 CCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTT 231
              |||   |   |||   | | |     || |||  | | |    | |    |||   | ||
Db        358 CCTTGGCTCTTTGAGCTCCCTTTTGCTTAATTACTGGGTTTTCTGGGCAGTTTTTTTTTT 417

Qy        232 AAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGAT 291
              |||    | | |  ||  || ||| | |||  |||    | ||  |||  |  || | |
Db        418 AAAGAGTTGGAGTAAGAAGAGGATTAAAAAGGAAAGGCNAATAANAAAAAAAAAAAAAAA 477

Qy        292 AAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGT 351
              ||||    |||| | |       |  || |    |  ||   ||   | || |||  |
Db        478 AAAAAAAAAAAAAAAAAAAAAAAAANAAAANNNAAAAAAAAAAAAAAANAAAAAAAAAAA 537

Qy        352 AAACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGAC 411
              ||    |||  | |   ||         | || || | | | |||   |       | |
Db        538 AANAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNNAAAA 597

Qy        412 TTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAAT 471
                ||| ||||                                  | |  |     ||||
Db        598 AAAAANAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNAAAAAAAAAANNNAAAAAN 657

Qy        472 ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGGTAAAAAACGC 531
                       |||     ||    ||||     |   ||| || ||  |   || |||
Db        658 NNNNNNNNNTTTTTTNAAAAAAAAAAANNNNAAAAAAAAAAAAAAANANNNAANAAAAAA 717

Qy        532 TTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAAT 591
                  ||    |  | || ||||   || ||   |||||| ||| ||  |   | |||||
Db        718 AANNAANNNAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAN 777

Qy        592 CAAAA 596
               ||||
Db        778 NAAAA 782

RESULT 40
CL037168/c
LOCUS       CL037168                1110 bp    DNA     linear   GSS 31-DEC-2003
DEFINITION  CH216-42N18_Sp6.1 CH216 Xenopus tropicalis genomic clone
            CH216-42N18, genomic survey sequence.
ACCESSION   CL037168
VERSION     CL037168.1  GI:40491901
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1110)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 14
            High quality sequence stop: 150.
FEATURES             Location/Qualifiers
     source          1..1110
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-42N18"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"
ORIGIN

  Query Match             7.8%;  Score 78.2;  DB 9;  Length 1110;
  Best Local Similarity   48.7%;  Pred. No. 1.8e-05;
  Matches  203;  Conservative    0;  Mismatches  214;  Indels    0;  Gaps    0;

Qy        180 TGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGT 239
              | || || || ||  || ||  ||| ||  |   ||  ||  || |  | ||     | |
Db        933 TTTTTTTTTTTTTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTAATTTTTTTTTTTTTT 874

Qy        240 TGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGG 299
              | |    |  |||  ||| ||||||||   |  |  |||  |  |  | | ||||
Db        873 TTGTTGTGAAGATAGTAAAAAAAAAAAANAAAAAAAAAAAAAAAAANNAAAAAAAAAAAA 814

Qy        300 AAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAA 359
              |||| | |       |  || |   ||  ||   ||   | || |||  |  |||   ||
Db        813 AAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAA 754

Qy        360 ACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAA 419
              |  | |   ||  |     || || || | | | |||   |   | ||| |   ||| ||
Db        753 AAAAAAAANAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 694

Qy        420 AATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGC 479
              ||        |  |   ||   || |      |   | |    ||  |||| |   |
Db        693 AAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAA 634

Qy        480 TTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGA 539
                   ||| ||    ||||     |   ||| || || ||   ||||||     | || |
Db        633 AAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 574

Qy        540 ACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAA 596
              | |  | || ||||   || || | ||||||| ||| ||  |   | |||||  ||||
Db        573 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 517

Search completed: August 26, 2005, 06:59:20
Job time : 3914 secs
```

15

```
                         GenCore version 5.1.6
              Copyright (c) 1993 - 2005  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         August 26, 2005, 05:54:23 ; Search time 2119 Seconds
                                           (without alignments)
                                           3112.522 Million cell updates/sec

Title:          US-10-724-972A-2580
Perfect score:  1008
Sequence:       1 ggagtggaatcagtgagagg..........gtaaaaaagataataagtaa 1008

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters:      14663426

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_NA:*
                1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
                10:  /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
                12:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
                13:  /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
                14:  /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
                15:  /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
                16:  /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
                17:  /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
                18:  /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
                19:  /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
                20:  /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
                21:  /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
                22:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
                23:  /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
                24:  /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
                25:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
                26:  /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                     %
Result               Query
  No.     Score     Match Length DB  ID                         Description
-----------------------------------------------------------------------------
    1      1008     100.0    1008 22  US-10-724-972A-2580         Sequence 2580, Ap
    2       993      98.5     993 17  US-10-282-122A-34954        Sequence 34954, A
    3       567      56.2     984 17  US-10-282-122A-7644         Sequence 7644, Ap
    4     564.4      56.0    1014  9  US-09-815-242-8156          Sequence 8156, Ap
    5     531.8      52.8     927  9  US-09-815-242-4404          Sequence 4404, Ap
    6     309.8      30.7     525 17  US-10-282-122A-35107        Sequence 35107, A
    7     265.6      26.3     960  9  US-09-815-242-8024          Sequence 8024, Ap
    8     264.6      26.2     957  9  US-09-815-242-4194          Sequence 4194, Ap
    9     263.2      26.1     957 21  US-10-470-048B-301          Sequence 301, App
   10     263.2      26.1     960 17  US-10-282-122A-7949         Sequence 7949, Ap
   11     222.6      22.1     668  8  US-08-781-986A-355          Sequence 355, App
   12     222.6      22.1     668 18  US-10-329-624-355           Sequence 355, App
   13     162.2      16.1     242  8  US-08-781-986A-2556         Sequence 2556, Ap
   14     162.2      16.1     242 18  US-10-329-624-2556          Sequence 2556, Ap
   15     156.4      15.5     897 17  US-10-282-122A-30706        Sequence 30706, A
   16     149.4      14.8     801  9  US-09-974-300-2066          Sequence 2066, Ap
   17     143.2      14.2     972 17  US-10-282-122A-9415         Sequence 9415, Ap
c  18       128      12.7     321  9  US-09-815-242-2107          Sequence 2107, Ap
c  19       128      12.7     321 17  US-10-282-122A-4605         Sequence 4605, Ap
   20     125.4      12.4    2115  8  US-08-781-986A-604          Sequence 604, App
   21     125.4      12.4    2115 18  US-10-329-624-604           Sequence 604, App
   22       124      12.3     990 21  US-10-470-048B-318          Sequence 318, App
   23       124      12.3     999 16  US-10-278-946-15            Sequence 15, Appl
   24       124      12.3     999 21  US-10-967-189-15            Sequence 15, Appl
   25       124      12.3    3775  8  US-08-781-986A-238          Sequence 238, App
   26       124      12.3    3775 18  US-10-329-624-238           Sequence 238, App
   27     113.2      11.2     796 17  US-10-282-122A-10009        Sequence 10009, A
   28     103.2      10.2     900 17  US-10-282-122A-19870        Sequence 19870, A
   29      95.2       9.4     909  9  US-09-815-242-6372          Sequence 6372, Ap
   30      95.2       9.4     909 17  US-10-282-122A-20621        Sequence 20621, A
c  31      95.2       9.4   10244 21  US-10-893-671-14            Sequence 14, Appl
   32        79       7.8     963 17  US-10-282-122A-17083        Sequence 17083, A
   33      78.8       7.8    1026 17  US-10-282-122A-32601        Sequence 32601, A
c  34        78       7.7     431  9  US-09-960-352-5558          Sequence 5558, Ap
   35      77.8       7.7     478 19  US-10-021-323-6774          Sequence 6774, Ap
   36      75.8       7.5     891 17  US-10-282-122A-33236        Sequence 33236, A
c  37      75.6       7.5    1121 20  US-10-425-115-115706        Sequence 115706,
c  38      74.2       7.4    1062 20  US-10-425-115-120013        Sequence 120013,
   39        74       7.3    1214 18  US-10-424-599-102083        Sequence 102083,
c  40        74       7.3    1243 20  US-10-425-115-172717        Sequence 172717,
c  41      73.8       7.3     421  8  US-08-781-986A-383          Sequence 383, App
c  42      73.8       7.3     421 18  US-10-329-624-383           Sequence 383, App
   43      72.8       7.2     945  9  US-09-738-626-375           Sequence 375, App
   44      72.8       7.2 3309400  9  US-09-738-626-1             Sequence 1, Appli
c  45      72.6       7.2     446  9  US-09-960-352-3400          Sequence 3400, Ap


                               ALIGNMENTS


RESULT 1
US-10-724-972A-2580
; Sequence 2580, Application US/10724972A
; Publication No. US20040147734A1
; GENERAL INFORMATION:
;  APPLICANT: Doucette-Stamm, Lynn
;  APPLICANT:  Bush, David
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: PATH03-16
;  CURRENT APPLICATION NUMBER: US/10/724,972A
;  CURRENT FILING DATE:  2003-12-01
;  PRIOR APPLICATION NUMBER: 09/450,969
;  PRIOR FILING DATE: 1999-11-29
;  PRIOR APPLICATION NUMBER: 09/134,001
;  PRIOR FILING DATE: 1998-08-13
;  PRIOR APPLICATION NUMBER: 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 7544
; SEQ ID NO 2580
;   LENGTH: 1008
;   TYPE: DNA
;   ORGANISM: S.epidermidis
US-10-724-972A-2580

  Query Match             100.0%;  Score 1008;  DB 22;  Length 1008;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-162;
  Matches 1008;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 GGAGTGGAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGAGTGGAATCAGTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTA 60

Qy         61 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAA 120

Qy        121 GATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGT 180

Qy        181 GTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTT 240

Qy        241 GGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGA 300

Qy        301 AAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAA 360

Qy        361 CCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAA 420

Qy        421 ATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCT 480

Qy        481 TTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAA 540

Qy        541 CACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTA 600

Qy        601 TTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGT 660

Qy        661 CAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTA 720

Qy        721 AGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAAT 780

Qy        781 CCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAA 840

Qy        841 CTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATT 900

Qy        901 TTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAA 960

Qy        961 GAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAGATAATAAGTAA 1008


RESULT 2
US-10-282-122A-34954
; Sequence 34954, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34954
;   LENGTH: 993
;   TYPE: DNA
;   ORGANISM: Staphylococcus epidermidis
US-10-282-122A-34954

  Query Match             98.5%;  Score 993;  DB 17;  Length 993;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-159;
  Matches  993;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 GTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCA 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCA 60

Qy         73 GCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAA 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAA 120

Qy        133 ATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTT 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTT 180

Qy        193 GAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGAT 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGAT 240

Qy        253 GATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCT 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCT 300

Qy        313 GTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATT 372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATT 360

Qy        373 ATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACG 432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACG 420

Qy        433 ATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATT 492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATT 480

Qy        493 TCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAA 552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAA 540

Qy        553 ATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTA 612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTA 600

Qy        613 GCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGT 672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGT 660

Qy        673 CAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTT 732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTT 720

Qy        733 AAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATG 792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATG 780

Qy        793 TTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGAT 852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGAT 840

Qy        853 CCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGAC 912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGAC 900

Qy        913 TTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAA 972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAA 960

Qy        973 TTATCTAAGAAAGATAGTAAAAAAGATAATAAG 1005
              |||||||||||||||||||||||||||||||||
Db        961 TTATCTAAGAAAGATAGTAAAAAAGATAATAAG 993


RESULT 3
US-10-282-122A-7644
; Sequence 7644, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7644
;   LENGTH: 984
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-10-282-122A-7644

  Query Match             56.2%;  Score 567;  DB 17;  Length 984;
  Best Local Similarity   73.6%;  Pred. No. 3.9e-87;
  Matches  723;  Conservative    0;  Mismatches  260;  Indels    0;  Gaps    0;

Qy         14 TGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAG 73
              |||||||| |||||| ||| ||| || | || ||| |  |||  ||| ||  |  |||||
Db          2 TGAGAGGTCTAAAAACTTTTAGTATATTGGGATTAATAGTTGCCTTACTTTTAGTTGCAG 61

Qy         74 CATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAA 133
              | ||||| ||||    |      |    ||   |  | ||| | ||||||    ||
Db         62 CTTGTGGTAATACGGATAATTCAAGTAAAAAAGAATCATCAACTAAAGATACTATTTCGG 121

Qy        134 TCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTG 193
              | ||  | ||| | |||||    ||||||||||||| |  | |||||| | |||||  | |
Db        122 TAAAAGATGAAAATGGTACAGTAAAAGTACCTAAAGATGCAAAACGTATCGTTGTATTAG 181

Qy        194 AGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATG 253
              |||| |||||||  ||||| ||||  || ||||| |||||||| ||||| || || ||||
Db        182 AGTACTCATTTGCAGATGCATTAGCAGCATTAGACGTTAAACCAGTTGGTATTGCTGATG 241

Qy        254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
              ||   || ||||| ||||| ||||||||| | ||||| ||||||||  | || |||||||
Db        242 ATGGTAAGAAAAAACGTATCATTAAACCAGTTAGAGAAAAAATTGGGGATTATACTTCTG 301

Qy        314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
              |||| |||||||| || || |||||||| ||||| |||||| | ||||| ||||||||||
Db        302 TAGGTACACGTAAACAGCCAAACTTAGAGGAAATTAGTAAATTAAAACCGGATTTAATTA 361

Qy        374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
              | |||||||  | |||||| ||||||||| ||||||| |||||| ||||||||| || ||
Db        362 TCGCTGATAGCAGTAGACATAAAGGTATTAATAAAGAATTAAACAAAATTGCACCAACAT 421

Qy        434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
              |   | | || ||||| |||||||| || ||  ||||||||| || | || |||||||||
Db        422 TATCATTAAAGAGTTTTGATGGAGACTACAAACAAAATATTAATTCGTTCAAAACAATTG 481

Qy        494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
              | |||||||||  ||||||| |||||||  |||||||  | |  || || |||||   ||
Db        482 CTAAAGCTTTAAATAAAGAAAAAGAAGGCGAAAAACGTCTTGCTGAGCATGATAAATTAA 541

Qy        554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAG 613
              |  |  | |||||| | ||||| |   | |||| ||||||||| ||  | || |||||||
Db        542 TCAATAAGTATAAAGATGAAATTAAATTTGATAGAAATCAAAAAGTGCTTCCAGCAGTAG 601

Qy        614 CTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTC 673
               ||||||| | |||||  | || ||||||| | | || |||||| |||||  |||
Db        602 TTGCTAAAGCTGGTTTATTAGCACATCCAAACTATTCATATGTTGGACAATTTTTAAACG 661

Qy        674 AACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTA 733
```

```
                ||||||| |||||| | |||||||||||| ||||| || |||||||||||||| ||| | |
Db           662 AACTAGGATTTAAAAATGCATTAAGTGACGATGTAACAAAAGGTTTAAGTAAATATTTGA 721

Qy           734 AAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGT 793
                 ||||||||||||||||| |  ||||||||  |||| || |  | ||||| ||||||||| 
Db           722 AAGGACCTTACTTACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGA 781

Qy           794 TCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATC 853
                 |||| |||||| |  | || |      |     ||  | | ||  |  || || ||||| 
Db           782 TCATTATGACAGATCATGCTAAAAAAGATTCTGCTGAATTCAAGAAGTTACAAGAAGATG 841

Qy           854 CTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACT 913
                 |   |||||| || || || || || |||||| | || || |||||| | ||||||||  
Db           842 CAACATGGAAAAAGTTGAATGCAGTTAAAAATAATCGCGTGGATATTGTTGACCGTGATG 901

Qy           914 TATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAAT 973
                 | ||||||||||| ||||| ||||||||||| ||||||||||| ||||||||||||||||
Db           902 TTTGGGCAAGATCTCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTTGTTGAAT 961

Qy           974 TATCTAAGAAAGATAGTAAAAAA 996
                 |||| || |||||    ||  ||
Db           962 TATCAAAAAAAGAACAAAAGTAA 984

RESULT 4
US-09-815-242-8156
; Sequence 8156, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8156
;   LENGTH: 1014
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(1014)
US-09-815-242-8156

  Query Match             56.0%;  Score 564.4;  DB 9;  Length 1014;
  Best Local Similarity   73.1%;  Pred. No. 1.1e-86;
  Matches  724;  Conservative    0;  Mismatches  266;  Indels    0;  Gaps    0;

Qy             7 GAATCAGTGAGAGGTTTAAAAAATTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCA 66
                 | | |  |||||||| |||||| ||| ||| || | || ||| |  |||  ||| ||  |
Db            25 GGAACGATGAGAGGTCTAAAAACTTTTAGTATATTGGGATTAATAGTTGCCTTATTTTTA 84

Qy            67 ACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGA 126
                   |||||| ||||| ||||    |      |    ||   ||   ||| | ||||||   
Db            85 GTTGCAGCTTGTGGTAATACGGATAATTCAAGTAAAAAAGAATCATCAACTAAAGATACT 144

Qy           127 GTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTT 186
                  ||    | ||  | ||| | |||||    |||||||||||| |  | |||||| | |||
Db           145 ATTTCGGTAAAAGATGAAAATGGTACAGTAAAAGTACCTAAAGATGCAAAACGTATCGTT 204

Qy           187 GTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATA 246
                 ||  | ||||| |||||||  ||||| ||||  || ||||| |||||||| ||||| || 
Db           205 GTATTAGAGTACTCATTTGCAGATGCATTAGCAGCATTAGACGTTAAACCAGTTGGTATT 264

Qy           247 GCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATAC 306
                 || ||||||   || ||||| ||||| ||||||||| | ||||| |||||||| || || 
Db           265 GCTGATGATGGTAAGAAAAAACGTATCATTAAACCAGTTAGAGAAAAAATTGGGAATTAT 324

Qy           307 ACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGAT 366
                 ||||||||||| |||||||| || || |||||||| ||||| |||||| | ||||| |||
Db           325 ACTTCTGTAGGTACACGTAAACAGCCAAACTTAGAGGAAATTAGTAAATTAAAACCGGAT 384

Qy           367 TTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCT 426
                 |||||||| |||||||  | |||||| ||||||||| ||||||| ||||| |||||||| 
Db           385 TTAATTATCGCTGATAGCAGTAGACATAAAGGTATTAATAAAGAATTAAACAAAATTGCA 444

Qy           427 CCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAA 486
                 || ||  |   | | || ||||| |||||||| || ||  |||| ||| || | || |||
Db           445 CCAACATTATCATTAAAGAGTTTTGATGGAGACTACAAACAAAACATTAATTCGTTCAAA 504

Qy           487 ACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGAT 546
                 |||||| | |||||||||  ||||||| |||||||  |||| ||  | |  ||||| |||
Db           505 ACAATTGCTAAAGCTTTAAATAAAGAAAAAGAAGGCGAAAAGCGTCTTGCTGAACATGAT 564

Qy           547 AAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCT 606
                 ||   |||  || | |||||| | ||||| |   | |||| ||||||||| ||  | || 
Db           565 AAATTAATCAAAAAGTATAAAGATGAAATTAAGTTTGATAGAAATCAAAAAGTGCTTCCA 624

Qy           607 GCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTC 666
                 ||||| | ||||||| | |||||  | || ||||||| | | || |||||||| ||||| 
Db           625 GCAGTTGTTGCTAAAGCTGGTTTATTAGCACATCCAAACTATTCATATGTTGGACAATTT 684

Qy           667 CTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAG 726
                  |||   |||| || |||||| | ||||||||||||||||| || |||||||||||||| 
Db           685 TTAAACGAACTTGGATTTAAAAATGCATTAAGTGATGATGTAACAAAAGGTTTAAGTAAA 744

Qy           727 TATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAG 786
                 ||  | |||||||||||||||||| |  | ||||||  |||| || |  | |||||||| 
Db           745 TACTTGAAAGGACCTTACTTACAATTAGATACTGAACATTTAGCTGACTTAAATCCTGAA 804

Qy           787 CGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAA 846
                 || ||| | || |||||| | || || |      |     ||  | | ||  |  || ||
Db           805 CGCATGATTATTATGACAGATAATGCTAAAAAAGATTCTGCTGAATTCAAGAAGTTACAA 864

Qy           847 AAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGAC 906
                  |||||||    ||||| || || ||||| || |||||| | || || |||||| | |||
Db           865 GAAGATCCAACTTGGAAAAAGTTGAACGCAGTTAAAAATAATCGCGTGGATATTGTTGAC 924

Qy           907 CGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTT 966
                 |||||  | ||||||||||| ||||| ||||||||||| ||||||||||| |||||||||
Db           925 CGTGATGTTTGGGCAAGATCTCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTT 984

Qy           967 GTTGAATTATCTAAGAAAGATAGTAAAAAA 996
                 ||||||||||| || |||||    ||  ||
Db           985 GTTGAATTATCAAAAAAAGAACAAAAGTAA 1014

RESULT 5
US-09-815-242-4404
; Sequence 4404, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4404
;   LENGTH: 927
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-09-815-242-4404

  Query Match             52.8%;  Score 531.8;  DB 9;  Length 927;
  Best Local Similarity   73.4%;  Pred. No. 3.8e-81;
  Matches  680;  Conservative    0;  Mismatches  247;  Indels    0;  Gaps    0;

Qy         46 TTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGT 105
              ||| |  |||  ||| ||  |  |||||| ||||| ||||    |      |    ||
Db          1 TTAATAGTTGCCTTATTTTTAGTTGCAGCTTGTGGTAATACGGATAATTCAAGTAAAAAA 60

Qy        106 AAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCT 165
               ||   ||| | ||||||    ||    | ||  | ||| | |||||    |||||||||
Db         61 GAATCATCAACTAAAGATACTATTTCGGTAAAAGATGAAAATGGTACAGTAAAAGTACCT 120

Qy        166 AAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTA 225
              ||| |  | |||||| | ||||||  | ||||| |||||||  ||||| ||||  || |||
Db        121 AAAGATGCAAAACGTATCGTTGTATTAGAGTACTCATTTGCAGATGCATTAGCAGCATTA 180

Qy        226 GATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTA 285
              || |||||||| ||||| || || ||||||   || ||||| ||||| ||||||||| |
Db        181 GACGTTAAACCAGTTGGTATTGCTGATGATGGTAAGAAAAAACGTATCATTAAACCAGTT 240

Qy        286 AGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAA 345
              ||||| |||||||| || || ||||||||||| |||||||| || || |||||| |||
Db        241 AGAGAAAAAATTGGGAATTATACTTCTGTAGGTACACGTAAACAGCCAAACTTAGAGGAA 300

Qy        346 ATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTAT 405
              || |||||| | ||||| ||||||||||| |||||||  | |||||| ||||||||| ||
Db        301 ATTAGTAAATTAAAACCGGATTTAATTATCGCTGATAGCAGTAGACATAAAGGTATTAAT 360

Qy        406 AAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAAT 465
              ||||| ||||| |||||||| || ||  |   | | || ||||| |||||||| || ||
Db        361 AAAGAATTAAACAAAATTGCACCAACATTATCATTAAAGAGTTTTGATGGAGACTACAAA 420

Qy        466 GAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAA 525
               |||| ||| || | || ||||||| | |||||||||  ||||||| ||||||||  ||
Db        421 CAAAACATTAATTCGTTCAAAACAATTGCTAAAGCTTTAAATAAAGAAAAAGAAGGCGAA 480

Qy        526 AAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGAT 585
              || ||  | |  ||||| |||||   |||  || | |||||| | ||||| |   | |||
Db        481 AAGCGTCTTGCTGAACATGATAAATTAATCAAAAAGTATAAAGATGAAATTAAGTTTGAT 540
```

```
Qy        586 AAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGC 645
              | ||||||||| ||  | || ||||| | |||||| | |||||  | || ||||||| |
Db        541 AGAAATCAAAAAGTGCTTCCAGCAGTTGTTGCTAAAGCTGGTTTATTAGCACATCCAAAC 600

Qy        646 AACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGAT 705
               | || |||||||| |||||  |||   ||||| || ||||||| | |||||||||||||||
Db        601 TATTCATATGTTGGACAATTTTTAAACGAACTTGGATTTAAAAATGCATTAAGTGATGAT 660

Qy        706 GTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACT 765
              || || |||||||||||||| ||  | |||||||||||||||||| |  | ||||||  |
Db        661 GTAACAAAAGGTTTAAGTAAATACTTGAAAGGACCTTACTTACAATTAGATACTGAACAT 720

Qy        766 TTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAA 825
              ||| || |  | |||||||| || ||| | || |||||| | || || |      |
Db        721 TTAGCTGACTTAAATCCTGAACGCATGATTATTATGACAGATAATGCTAAAAAAGATTCT 780

Qy        826 CCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAAT 885
               ||  | | ||  |  || || |||||||    ||||| || || ||||| || ||||||
Db        781 GCTGAATTCAAGAAGTTACAAGAAGATCCAACTTGGAAAAAGTTGAACGCAGTTAAAAAT 840

Qy        886 CAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCA 945
               | || || |||||| | ||||||||  | ||||||||||| ||||| |||||||||||
Db        841 AATCGCGTGGATATTGTTGACCGTGATGTTTGGGCAAGATCTCGTGGCTTAATTTCTTCT 900

Qy        946 GAAGAAATGGCAAAAGAACTTGTTGAA 972
              ||||||||||| |||||||||||||||
Db        901 GAAGAAATGGCTAAAGAACTTGTTGAA 927


RESULT 6
US-10-282-122A-35107
; Sequence 35107, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35107
;   LENGTH: 525
;   TYPE: DNA
;   ORGANISM: Staphylococcus haemolyticus
US-10-282-122A-35107

  Query Match             30.7%;  Score 309.8;  DB 17;  Length 525;
  Best Local Similarity   76.3%;  Pred. No. 2.1e-43;
  Matches  399;  Conservative    0;  Mismatches  112;  Indels   12;  Gaps    1;

Qy          14 TGAGAGGTTTAAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAG 73
               ||| ||| || ||  ||  | || || | |  ||||| ||||  ||| ||  |  | | |
Db           2 TGAAAGGCTTTAAGTTTGCAGGTATAGTCGCATTATTATTTGCATTAGTTTTAGTTACTG 61

Qy          74 CATGTGGAAATAATAGTTCAAGTAACTCAAG------------TAAAGAGTCATCAAAAG 121
               | ||||| |||  |||    | |   ||| |            ||||  |||||| ||||
Db          62 CTTGTGGTAATGTTAGCAATAATGGATCAGGTGATTCAGGTAATAAATCGTCATCGAAAG 121

Qy         122 ATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTG 181
               ||    || |||| || ||||||  ||| || || ||||| |||||| |  |||||||||
Db         122 ATTCTATTAAAATTAAACACGAATTAGGCACAACTAAAGTTCCTAAAGATGCTAAACGTG 181

Qy         182 TTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTG 241
               | || |   | || | ||||||||| ||||| ||||  ||  || |||||||||||||||
Db         182 TCGTAGCGTTGGAATTTTCATTTGTAGATGCTTTAGCGGCACTAAATGTTAAACCTGTTG 241

Qy         242 GGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAA 301
               || | || ||||| ||||||  |||||||||||||||||||||||| ||| ||||||||| 
Db         242 GGGTTGCTGATGACAACAAACCAAATCGTATTATTAAACCATTAAAAGAAAAAATTGGAG 301

Qy         302 AATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAAC 361
               | || |  ||||| ||  | ||||| ||||| |||||||||||||||||||||| | ||||
Db         302 ATTATAAATCTGTTGGTGCTCGTAAACAACCCAACTTAGAAGAAATCAGTAAATTAAAAC 361

Qy         362 CAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAA 421
               |||||||||||||||| |||| ||| ||||||||||||||||| ||||| |||| |||||
Db         362 CAGATTTAATTATTGCAGATAGTAACAGACACAAAGGTATTTACAAAGAATTAAGTAAAA 421

Qy         422 TTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTT 481
               ||||||| || |||||| | |||||||| ||||||||||||||||| ||||| |||||||
Db         422 TTGCTCCAACAATTGAATTAAAAAGTTTTGATGGAGATTATAATGATAATATAGATGCTT 481

Qy         482 TTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAA 524
               ||||||||||| | |||||||||  |||||| || | |||| |
Db         482 TTAAAACAATTGCCAAAGCTTTAAATAAAGATGATGTAGGTCA 524

RESULT 7
US-09-815-242-8024
; Sequence 8024, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
```

```
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8024
;   LENGTH: 960
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(960)
US-09-815-242-8024

  Query Match             26.3%;  Score 265.6;  DB 9;  Length 960;
  Best Local Similarity   57.8%;  Pred. No. 8.2e-36;
  Matches  494;  Conservative    0;  Mismatches  354;  Indels    6;  Gaps    1;

Qy         132 AATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCT 191
               ||| ||  | |||  ||| |||   ||| |    ||| | |||||||||||||||||  |
Db         105 AATAAAAGATGAATTAGGAACTGAAAAAATTAAGAAAAATCCTAAACGTGTTGTTGTATT 164

Qy         192 TGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGA 251
                || |||   |||| ||||   ||||  || |||||| | ||||||||||| || || ||
Db         165 AGAATATAGTTTTGCTGATTATTTAGCAGCATTAGATATGAAACCTGTTGGTATTGCAGA 224

Qy         252 TGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTC 311
               ||||  ||  |  ||   ||| |  ||  || ||||||||||  ||||   |||    ||
Db         225 TGATGGCAGCAGTAAAAATATAACAAAGTCAGTAAGAGATAAGGTTGGGGCATATGAATC 284

Qy         312 TGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAAT 371
                || ||| |  |   ||||||| ||  | ||||  || |||||| | ||||| ||||| ||
Db         285 GGTTGGATCTAGACCGCAACCGAATATGGAAGTGATAAGTAAATTAAAACCGGATTTGAT 344

Qy         372 TATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTAC 431
                || || |||  ||  ||||| ||    ||  |   ||| ||||  ||||||||||| ||
Db         345 CATAGCAGATGTTAGCAGACATAAGAAAATCAAATCAGAATTAAGCAAAATTGCTCCGAC 404

Qy         432 GATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAAT 491
                ||     |    ||       ||||||||||||| |||||||||||| |||||||| |
Db         405 AATCATGTTAGTTAGCGGTACGGGAGATTATAATGCAAATATTGATGCATTTAAAACAGT 464

Qy         492 TTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAA 551
                 | |||||  |||| |||||  |||||||  | |||||  | ||| | || ||||| | 
Db         465 CGCTAAAGCAGTAGGCAAAGAGAAAGAAGGCGAGAAACGTCTGGAAAAGCATGATAAAAT 524

Qy         552 AATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGT 611
               | | |  ||   || |||  ||||      |  ||      |||  | |||  |    | 
Db         525 ATTAGCGGAGATTAGAAAGAAAATTGAACAGAGTACGTTAAAAACTGCATTTGCATTCGG 584

Qy         612 AGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAG 671
                    | | | ||||| || ||  | ||       |  | | | || |||||| ||| 
Db         585 TATCTCAAGAGCAGGTATGTTTATTAATAATGAAGATACATTTATGGGACAATTCTTAAT 644

Qy         672 TCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCT 731
               | || | ||| || ||   | | | |   | ||    ||||    | |  || |      
Db         645 TAAAATGGGTATTCAACCTGAAGTCACAAAAGACAAAACTACGCATGTTGGTGAACGCAA 704

Qy         732 TAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTAT 791
                   || |||||  |  |  | || | ||||    |  |  |  | |||||  |   |||
Db         705 GGGTGGTCCTTATATTTATTTAAATAATGAAGAACTTGCCAATATCAATCCAAAAGTTAT 764

Qy         792 GTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGA 851
               | |  ||   ||  ||  |  ||     ||  |         ||   || |     |   
Db         765 GATTTTAGCCACTGACGGAAAAACGGACAAAAATAG------AACGAAATTCATTGATCC 818

Qy         852 TCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGA 911
               | | || ||||||   ||||||| ||||||||| || |    ||| ||    | |||||  |
Db         819 TGCAGTTTGGAAATCATTAAAAGCTGTGAAAGATAACAAAGTTTATGACGTTGACCGAAA 878
```

```
Qy         912 CTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGA 971
                   |||   | |||| | ||| | ||  |    |||   |||||| ||||  | |   |
Db         879 TAAGTGGTTGAAATCAAGGGGTATTATCGCAAGTGAAAGTATGGCAGAAGATTTAGAAAA 938

Qy         972 ATTATCTAAGAAAG 985
               | |  |  | ||||
Db         939 AATTGCAGAAAAAG 952


RESULT 8
US-09-815-242-4194
; Sequence 4194, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4194
;   LENGTH: 957
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-09-815-242-4194

  Query Match             26.2%;  Score 264.6;  DB 9;  Length 957;
  Best Local Similarity   57.8%;  Pred. No. 1.2e-35;
  Matches  493;  Conservative    0;  Mismatches  354;  Indels    6;  Gaps    1;

Qy         132 AATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCT 191
               ||| ||  | |||  ||| |||   ||| |    ||| | |||||||||||||||||  |
Db         111 AATAAAAGATGAATTAGGAACTGAAAAAATTAAGAAAAATCCTAAACGTGTTGTTGTATT 170

Qy         192 TGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGA 251
                || |||   |||| ||||   ||||  || |||||| | ||||||||||| || || ||
Db         171 AGAATATAGTTTTGCTGATTATTTAGCAGCATTAGATATGAAACCTGTTGGTATTGCAGA 230

Qy         252 TGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTC 311
               ||||  ||  |  ||   ||| |  ||  || ||||||||||  ||||   |||    ||
Db         231 TGATGGCAGCAGTAAAAATATAACAAAGTCAGTAAGAGATAAGGTTGGGGCATATGAATC 290

Qy         312 TGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAAT 371
                || ||| |  |   |||||| ||  | ||||  || |||||| | ||||| ||||| ||
Db         291 GGTTGGATCTAGACCGCAACCGAATATGGAAGTGATAAGTAAATTAAAACCGGATTTGAT 350

Qy         372 TATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTAC 431
                || || |||  ||  ||||| ||    ||   |   ||| ||||  ||||||||||| ||
Db         351 CATAGCAGATGTTAGCAGACATAAGAAAATCAAATCAGAATTAAGCAAAATTGCTCCGAC 410
```

```
Qy         432 GATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAAT 491
                ||     |    ||       ||||||||||||| |||||||||||| |||||||| |
Db         411 AATCATGTTAGTTAGCGGTACGGGAGATTATAATGCAAATATTGATGCATTTAAAACAGT 470

Qy         492 TTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAA 551
                 | |||||  |||| |||||  |||||||  | |||||  | ||| | || ||||| |
Db         471 CGCTAAAGCAGTAGGCAAAGAGAAAGAAGGCGAGAAACGTCTGGAAAAGCATGATAAAAT 530

Qy         552 AATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGT 611
               | | |  ||   || |||  ||||      |  ||      |||  | |||  |    |
Db         531 ATTAGCGGAGATTAGAAAGAAAATTGAACAGAGTACGTTAAAAACTGCATTTGCATTCGG 590

Qy         612 AGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAG 671
                    | | | ||||| || ||  | ||       |  | | | | || ||||||| |||
Db         591 TATCTCAAGAGCAGGTATGTTTATTAATAATGAAGATACATTTATGGGACAATTCTTAAT 650

Qy         672 TCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCT 731
               | || | ||| || ||   | | | |   | ||    ||||    | |  || |
Db         651 TAAAATGGGTATTCAACCTGAAGTCACAAAAGACAAAACTACGCATGTTGGTGAACGCAA 710

Qy         732 TAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTAT 791
                   || |||||  |  |  | || | ||||    |  |  |  | |||||  |   |||
Db         711 GGGTGGTCCTTATATTTATTTAAATAATGAAGAACTTGCCAATATCAATCCAAAAGTTAT 770

Qy         792 GTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGA 851
               | |  ||   ||  ||  |  ||     ||  |         ||   || |     |
Db         771 GATTTTAGCCACTGACGGAAAAACGGACAAAAATAG------AACGAAATTCATTGATCC 824

Qy         852 TCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGA 911
               | | || |||||   |||||| |||||||||| || |    ||| ||    | |||||  |
Db         825 TGCAGTTTGGAAATCATTAAAAGCTGTGAAAGATAACAAAGTTTATGACGTTGACCGAAA 884

Qy         912 CTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGA 971
                   |||   | |||| | ||| | ||  |    |||   |||||| ||||  | |   |
Db         885 TAAGTGGTTGAAATCAAGGGGTATTATCGCAAGTGAAAGTATGGCAGAAGATTTAGAAAA 944

Qy         972 ATTATCTAAGAAA 984
               | |  |  | |||
Db         945 AATTGCAGAAAAA 957


RESULT 9
US-10-470-048B-301
; Sequence 301, Application US/10470048B
; Publication No. US20050037444A1
; GENERAL INFORMATION:
;  APPLICANT: MEINKE ET AL.
;  TITLE OF INVENTION: A METHOD FOR IDENTIFICATION, ISOLATION AND PRODUCTION OF
;  TITLE OF INVENTION:  ANTIGENS TO A SPECIFIC PATHOGEN
;  FILE REFERENCE: SONN:035US
;  CURRENT APPLICATION NUMBER: US/10/470,048B
;  CURRENT FILING DATE:  2003-07-25
;  NUMBER OF SEQ ID NOS: 603
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 301
;   LENGTH: 957
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-10-470-048B-301

  Query Match             26.1%;  Score 263.2;  DB 21;  Length 957;
  Best Local Similarity   55.9%;  Pred. No. 2.1e-35;
  Matches  524;  Conservative    0;  Mismatches  408;  Indels    6;  Gaps    1;

Qy          48 ATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAA 107
               | |  |||| || ||   || |  ||     | |       |||   || |   |  ||
Db          21 ACTAGTTGTGTTCATGCTAATTTTAGTTGTAGCAGTAGCGGGTTGTGGTCAAAAAGATAC 80

Qy         108 AGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAA 167
                ||      ||  ||         ||| ||  | |||  |||  || |||   ||| |    ||
Db          81 TGAAGAGAAAACTGAAATGACGACAATAAAAGATGAATTAGGAACTGAAAAAATTAAGAA 140
```

```
Qy        168 ACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGA 227
              | | |||||||||||||||||  | || |||   |||| ||||   ||||  || |||||
Db        141 AAATCCTAAACGTGTTGTTGTATTAGAATATAGTTTTGCTGATTATTTAGCAGCATTAGA 200

Qy        228 TGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAG 287
              | | ||||||||||| || || ||||||  ||  |  ||   ||| |  ||  || ||||
Db        201 TATGAAACCTGTTGGTATTGCAGATGATGGCAGCACTAAAAATATAACAAAGTCAGTAAG 260

Qy        288 AGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAAT 347
              |||||| |||||    |||   || || ||| |  |   |||||| ||  | ||||  ||
Db        261 AGATAAGATTGGGGCATATGAATCGGTTGGATCTAGACCGCAACCGAATATGGAAGTGAT 320

Qy        348 CAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAA 407
               |||||| | ||||| ||||| || ||||| |||  ||  ||||| ||    ||  |
Db        321 AAGTAAATTAAAACCGGATTTGATCATTGCAGATGTTAGCAGACATAAGAAAATCAAATC 380

Qy        408 AGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGA 467
              ||| || |  ||||||||||| || ||     |    ||       |||||||||||||
Db        381 AGAATTGAGCAAAATTGCTCCGACAATCATGTTAGTTAGCGGTACGGGAGATTATAATGC 440

Qy        468 AAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAA 527
              ||||||||| || ||||||||| |  | |||||  |||| |||||  |||||||  | ||
Db        441 AAATATTGAAGCATTTAAAACAGTCGCTAAAGCAGTAGGCAAAGAGAAAGAAGGCGAGAA 500

Qy        528 ACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAA 587
               ||  | ||| | || ||||| | | | |  ||   || |||  ||||      |  ||
Db        501 GCGTCTGGAAAAGCATGATAAAATATTAGCGGAGATTAGAAAGAAAATTGAACAGAGTAC 560

Qy        588 AAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAA 647
                   ||   | |||  |    |      | | | ||||| || ||  | ||       |
Db        561 GTTAAAATCTGCATTTGCATTCGGTATCTCAAGAGCAGGTATGTTTATTAATAATGAAGA 620

Qy        648 CTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGT 707
                | | | | || |||||| ||| | || | ||| || ||   | | | |   | ||
Db        621 TACATTTATGGGACAATTCTTAATTAAAATGGGTATTCAACCTGAAGTCACAAAAGACAA 680

Qy        708 TACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTT 767
               ||||    | |  || |          || |||||  || |  | || | ||||    |
Db        681 AACTACGCATGTTGGTGAACGCAAGGGTGGTCCTTATATATATTTAAATAATGAAGAACT 740

Qy        768 ATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACC 827
                |  |  | |||||  |   |||| |  ||   ||  ||  |  ||     ||  |
Db        741 TGCCAATATCAATCCAAAAGTTATGATTTTAGCCACTGACGGAAAAACGGACAAAAA--- 797

Qy        828 TTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCA 887
                    ||   || |     |   | | || |||||   |||||| ||||||||| || |
Db        798 ---TAGAACGAAATTCATTGATCCTGCAGTTTGGAAATCATTAAAAGCTGTGAAAGATAA 854

Qy        888 ACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGA 947
                  ||| ||    | |||||  |    |||   | |||| | ||  | ||  |    ||
Db        855 CAAAGTTTATGACGTTGACCGAAATAAGTGGTTGAAATCAAGGGGGATTATCGCAAGTGA 914

Qy        948 AGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAG 985
              |   |||||| ||||  | |   || |  |  | ||||
Db        915 AAGTATGGCAGAAGATTTAGAAAAAATTGCAGAAAAAG 952

RESULT 10
US-10-282-122A-7949
; Sequence 7949, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7949
;   LENGTH: 960
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-10-282-122A-7949

  Query Match             26.1%;  Score 263.2;  DB 17;  Length 960;
  Best Local Similarity   55.9%;  Pred. No. 2.1e-35;
  Matches  524;  Conservative    0;  Mismatches  408;  Indels    6;  Gaps    1;

Qy         48 ATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAATAGTTCAAGTAACTCAAGTAA 107
              |||  |||| || ||   ||    ||     | |       |||   || |   |  ||
Db         21 ATTAGTTGTGTTTATGCTAATCTTAGTTGTAGCAGTAGCGGGTTGTGGTCAAAAAGATAC 80

Qy        108 AGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAA 167
               ||      ||  ||         ||| ||  | |||  ||| |||   ||| |    ||
Db         81 TGAAGAGAAAACTGAAATGACGACAATAAAAGATGAATTAGGAACTGAAAAAATTAAGAA 140

Qy        168 ACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGA 227
              | | |||||||||||||||||  | || |||   |||| ||||   ||||  || |||||
Db        141 AAATCCTAAACGTGTTGTTGTATTAGAATATAGTTTTGCTGATTATTTAGCAGCATTAGA 200

Qy        228 TGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAG 287
              | | ||||||||||| || || ||||||  ||  |  ||   ||| |  ||  || ||||
Db        201 TATGAAACCTGTTGGTATTGCAGATGATGGCAGCACTAAAAATATAACAAAGTCAGTAAG 260

Qy        288 AGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAAT 347
              |||||| |||||    |||   || || ||| |  |   |||||| ||  | ||||  ||
Db        261 AGATAAGATTGGGGCATATGAATCGGTTGGATCTAGACCGCAACCGAATATGGAAGTGAT 320

Qy        348 CAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAA 407
               |||||| | ||||| ||||| || ||||| |||  ||  ||||| ||    ||  |
Db        321 AAGTAAATTAAAACCGGATTTGATCATTGCAGATGTTAGCAGACATAAGAAAATCAAATC 380

Qy        408 AGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGA 467
              ||| || |  ||||||||||| || ||     |    ||       |||||||||||||
Db        381 AGAATTGAGCAAAATTGCTCCGACAATCATGTTAGTTAGCGGTACGGGAGATTATAATGC 440

Qy        468 AAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAA 527
              ||||||||| || ||||||||| |  | |||||  |||| |||||  |||||||  | ||
Db        441 AAATATTGAAGCATTTAAAACAGTCGCTAAAGCAGTAGGCAAAGAGAAAGAAGGCGAGAA 500

Qy        528 ACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAA 587
               ||  | ||| | || ||||| | | | |  ||   || |||  ||||      |  ||
Db        501 GCGTCTGGAAAAGCATGATAAAATATTAGCGGAGATTAGAAAGAAAATTGAACAGAGTAC 560
```

```
Qy        588 AAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAA 647
                   ||   | |||  |    |      | | | ||||| || ||  | ||       |
Db        561 GTTAAAATCTGCATTTGCATTCGGTATCTCAAGAGCAGGTATGTTTATTAATAATGAAGA 620

Qy        648 CTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGT 707
                | | | | || |||||| ||| | || | ||| || ||   | | | |   | ||
Db        621 TACATTTATGGGACAATTCTTAATTAAAATGGGTATTCAACCTGAAGTCACAAAAGACAA 680

Qy        708 TACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTT 767
               ||||    | |  || |          || |||||  || |  | || | ||||    |
Db        681 AACTACGCATGTTGGTGAACGCAAGGGTGGTCCTTATATATATTTAAATAATGAAGAACT 740

Qy        768 ATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACC 827
                |  |  | |||||  |   |||| |  ||   ||  ||  |  ||     ||  |
Db        741 TGCCAATATCAATCCAAAAGTTATGATTTTAGCCACTGACGGAAAAACGGACAAAAA--- 797

Qy        828 TTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCA 887
                    ||   || |     |   | | || |||||   |||||| ||||||||| || |
Db        798 ---TAGAACGAAATTCATTGATCCTGCAGTTTGGAAATCATTAAAAGCTGTGAAAGATAA 854

Qy        888 ACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGA 947
                  ||| ||    | |||||  |    |||   | |||| | ||  | |  |  |    ||
Db        855 CAAAGTTTATGACGTTGACCGAAATAAGTGGTTGAAATCAAGGGGGATTATCGCAAGTGA 914

Qy        948 AGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAG 985
              |   |||||| ||||  | |   || |  |  | ||||
Db        915 AAGTATGGCAGAAGATTTAGAAAAAATTGCAGAAAAAG 952
```

```
RESULT 11
US-08-781-986A-355
; Sequence 355, Application US/08781986A
; Publication No. US20030054436A1
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  5255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,986A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Bob
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PB248PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  355:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  668 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-781-986A-355

  Query Match             22.1%;  Score 222.6;  DB 8;  Length 668;
```

```
  Best Local Similarity   74.8%;  Pred. No. 1.6e-28;
  Matches  279;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy        627 TTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAA 686
              |||  | || ||||||| | | || |||||||| ||||| |||   ||||||| |||||
Db          1 TTTATTAGCACATCCAAACTATTCATATGTTGGACAATTTTTAAACGAACTAGGATTTAA 60

Qy        687 AGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTT 746
              | | |||||||||| |||||| || ||||||||||||| ||| | ||||||||||||||
Db         61 AAATGCATTAAGTGACGATGTAACAAAAGGTTTAAGTAAATATTTGAAAGGACCTTACTT 120

Qy        747 ACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAA 806
              |||| |  ||||||||  |||| || |  | ||||| ||||||||| |||| |||||| |
Db        121 ACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGATCATTATGACAGA 180

Qy        807 CAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAA 866
                | ||  |     |     ||  | | ||  |  || || ||||| |   |||||| ||
Db        181 TCATGCTAAAAAAGATTCTGCTGAATTCAAGAAGTTACAAGAAGATGCAACATGGAAAAA 240

Qy        867 ATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATC 926
               || || || || |||||| | || || ||||| | ||||||||  | | ||||||||||
Db        241 GTTGAATGCAGTTAAAAATAATCGCGTGGATATTGTTGACCGTGATGTTTGGGCAAGATC 300

Qy        927 ACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGA 986
               ||||| |||||||||| ||||||||||| |||||||||||||||||||| || |||||
Db        301 TCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTTGTTGAATTATCAAAAAAAGA 360

Qy        987 TAGTAAAAAAGAT 999
                  ||  ||| |
Db        361 ACAAAAGTAAGGT 373
```

```
RESULT 12
US-10-329-624-355
; Sequence 355, Application US/10329624
; Publication No. US20040043037A1
;  GENERAL INFORMATION:
;       APPLICANT: Charles Kunsch
;                  Gil H. Choi
;                  Patrick S. Dillon
;                  Craig A. Rosen
;                  Steven C. Barash
;                  Michael R. Fannon
;       TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;       NUMBER OF SEQUENCES: 5256
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Human Genome Sciences, Inc.
;            STREET: 9410 Key West Avenue
;            CITY: Rockville
;            STATE: Maryland
;            COUNTRY: USA
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;            COMPUTER: HP Vectra 486/33
;            OPERATING SYSTEM: MSDOS version 6.2
;            SOFTWARE: ASCII Text
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/329,624
;            FILING DATE: 27-Dec-2002
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/956,171
;            FILING DATE: October 20, 1997
;            APPLICATION NUMBER: 60/009,861
;            FILING DATE: January 5, 1996
;            APPLICATION NUMBER: 08/781,986
;            FILING DATE: January 3, 1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Mark J. Hyman
;            REGISTRATION NUMBER: 46,789
;            REFERENCE/DOCKET NUMBER: PB248P1D1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (240) 314-1224
;             TELEFAX: (301) 309-8439
;    INFORMATION FOR SEQ ID NO: 355:
;         SEQUENCE CHARACTERISTICS:
;             LENGTH: 668 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;         SEQUENCE DESCRIPTION: SEQ ID NO: 355:
US-10-329-624-355

  Query Match             22.1%;  Score 222.6;  DB 18;  Length 668;
  Best Local Similarity   74.8%;  Pred. No. 1.6e-28;
  Matches  279;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy        627 TTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAA 686
              |||  | || ||||||| | | || |||||||| |||||  |||   ||||||| |||||
Db          1 TTTATTAGCACATCCAAACTATTCATATGTTGGACAATTTTTAAACGAACTAGGATTTAA 60

Qy        687 AGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTT 746
              | | ||||||||||| |||||| || ||||||||||||||| ||| | ||||||||||||||
Db         61 AAATGCATTAAGTGACGATGTAACAAAAGGTTTAAGTAAATATTTGAAAGGACCTTACTT 120

Qy        747 ACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAA 806
              |||| |  ||||||||  |||| || |  | ||||| ||||||||| |||| |||||| |
Db        121 ACAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGATCATTATGACAGA 180

Qy        807 CAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAA 866
                | || |      |     ||  | | ||  |  || || ||||| |   |||||| ||
Db        181 TCATGCTAAAAAAGATTCTGCTGAATTCAAGAAGTTACAAGAAGATGCAACATGGAAAAA 240

Qy        867 ATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATC 926
               || || || || |||||| | || || |||| | |||||||  | |||||||||||
Db        241 GTTGAATGCAGTTAAAAATAATCGCGTGGATATTGTTGACCGTGATGTTTGGGCAAGATC 300

Qy        927 ACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGA 986
               ||||| |||||||||| ||||||||||| |||||||||||||||||||||| || |||||
Db        301 TCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAAGAACTTGTTGAATTATCAAAAAAAGA 360

Qy        987 TAGTAAAAAAGAT 999
                  ||  ||| |
Db        361 ACAAAAGTAAGGT 373

RESULT 13
US-08-781-986A-2556
; Sequence 2556, Application US/08781986A
; Publication No. US20030054436A1
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  5255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,986A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Bob
;      REGISTRATION NUMBER:  30,446
```

```
;      REFERENCE/DOCKET NUMBER:  PB248PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  2556:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  242 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-781-986A-2556

  Query Match             16.1%;  Score 162.2;  DB 8;  Length 242;
  Best Local Similarity   79.3%;  Pred. No. 2.3e-18;
  Matches  191;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy        208 GATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAAT 267
              ||||| ||||  || ||||| |||||||| ||||| || | ||||||    || |||||
Db          2 GATGCATTAGCAGCATTAGACGTTAAACCAGTTGGTATTGCTGATGATGGTAAGAAAAAA 61

Qy        268 CGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAG 327
              ||||| ||||||||| | || ||| ||||||||  | || |||||||||| ||||||||
Db         62 CGTATCATTAAACCAGTTAGAGAAAAAATTGGGGATTATACTTCTGTAGGTACACGTAAA 121

Qy        328 CAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAAT 387
              || || |||||||||||||| |||||| | ||||| ||||||||||| |||||||  | |
Db        122 CAGCCAAACTTAGAAGAAATTAGTAAATTAAAACCGGATTTAATTATCGCTGATAGCAGT 181

Qy        388 AGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGT 447
              ||||| | |||||| |||||||| ||||  |||||||| || ||  |   | | || |||
Db        182 AGACATANAGGTATTAATAAAGAATTAANCAAAATTGCACCAACATTATCATTAAAGAGT 241

Qy        448 T 448
              |
Db        242 T 242

RESULT 14
US-10-329-624-2556
; Sequence 2556, Application US/10329624
; Publication No. US20040043037A1
;   GENERAL INFORMATION:
;        APPLICANT: Charles Kunsch
;                   Gil H. Choi
;                   Patrick S. Dillon
;                   Craig A. Rosen
;                   Steven C. Barash
;                   Michael R. Fannon
;        TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;        NUMBER OF SEQUENCES: 5256
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/329,624
;             FILING DATE: 27-Dec-2002
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/956,171
;             FILING DATE: October 20, 1997
;             APPLICATION NUMBER: 60/009,861
;             FILING DATE: January 5, 1996
;             APPLICATION NUMBER: 08/781,986
;             FILING DATE: January 3, 1997
;        ATTORNEY/AGENT INFORMATION:
```

```
;           NAME: Mark J. Hyman
;           REGISTRATION NUMBER: 46,789
;           REFERENCE/DOCKET NUMBER: PB248P1D1
;       TELECOMMUNICATION INFORMATION:
;           TELEPHONE: (240) 314-1224
;           TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 2556:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 242 base pairs
;           TYPE: nucleic acid
;           STRANDEDNESS: double
;           TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2556:
US-10-329-624-2556

  Query Match             16.1%;  Score 162.2;  DB 18;  Length 242;
  Best Local Similarity   79.3%;  Pred. No. 2.3e-18;
  Matches  191;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy        208 GATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAAT 267
              ||||| ||||  || ||||| |||||||| ||||| || || ||||||   || |||||
Db          2 GATGCATTAGCAGCATTAGACGTTAAACCAGTTGGTATTGCTGATGATGGTAAGAAAAAA 61

Qy        268 CGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAAG 327
              ||||| ||||||||| | ||||| ||||||||  | || ||||||||||| ||||||||
Db         62 CGTATCATTAAACCAGTTAGAGAAAAAATTGGGGATTATACTTCTGTAGGTACACGTAAA 121

Qy        328 CAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAAT 387
              || || |||||||||||||| |||||| | ||||| ||||||||||| |||||||  | |
Db        122 CAGCCAAACTTAGAAGAAATTAGTAAATTAAAACCGGATTTAATTATCGCTGATAGCAGT 181

Qy        388 AGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAGT 447
              ||||| | ||||||| ||||||| ||||  |||||||| || ||  |   | | || |||
Db        182 AGACATANAGGTATTAATAAAGAATTAANCAAAATTGCACCAACATTATCATTAAAGAGT 241

Qy        448 T 448
              |
Db        242 T 242


RESULT 15
US-10-282-122A-30706
; Sequence 30706, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30706
;   LENGTH: 897
;   TYPE: DNA
;   ORGANISM: Pasteurella multocida
US-10-282-122A-30706

  Query Match             15.5%;  Score 156.4;  DB 17;  Length 897;
  Best Local Similarity   51.9%;  Pred. No. 3.2e-17;
  Matches  454;  Conservative    0;  Mismatches  381;  Indels   39;  Gaps    3;

Qy        123 TGGAGTTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGT 182
              ||  |||    ||||  |  || ||||       |   ||  ||     |||||||||||
Db         63 TGCTGTTACCGTCAAAGATCAAAAAGGGGAATTTACCTTAGATAGCGTACCTAAACGTGT 122

Qy        183 TGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGG 242
              |||||   | || ||||| | ||| |||||  | |      | | ||| |  ||||||||
Db        123 TGTTGCCTTAGAATATTCTTATGTGGATGCACTTGCACAAATTGGTGTCAGCCCTGTTGG 182

Qy        243 GATAGCGGATGATAACAAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAA 302
                | || ||||| ||  | ||||  |||||| |  ||  | || | |||||| |
Db        183 TGTCGCAGATGACAATGATAAAACCCGTATTCTGCAAAAAGTACGCGATAAAGTGCAGCC 242

Qy        303 ATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACC 362
              ||     || || || || |||   ||||| |  |||||||  ||   |  |||||||||
Db        243 ATGGGAATCAGTGGGGACGCGTTCTCAACCGAGTTTAGAAGCGATTTCTGCACTTAAACC 302

Qy        363 AGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAAT 422
              |||||| || ||||| ||| ||||| | ||    |   | ||| ||||  | || |||||
Db        303 AGATTTGATCATTGCCGATGATAATCGCCATTCTGCCGTCTATGAAGAACTCAAAAAAAT 362

Qy        423 TGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTT 482
               || || ||  | |   | ||      | |||   | ||| | |||||  ||||  |
Db        363 CGCGCCGACAGTCGTCTTTAATTCCCGCCATGAGAACTATCAAGAAAACCTTGAAACCGC 422

Qy        483 TAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACA 542
              ||| |||    |    ||||||||| | |||||    ||   ||  |  |  |  ||||
Db        423 ACAAAAAATCGGTGATTTATTAGGTAAATCAAAAGAAATGCAAGCGCGTATTGCAAAACA 482

Qy        543 CGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATT 602
                |  || | ||||  |      | |  |  | |||  |    |||    |||| |   |
Db        483 TAAGCAGGATATTGCGG------ACATCGCCAAAACGTTACCGAAAGGGAAAAAAGCGAT 536

Qy        603 GCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCA 662
                  |  ||  |   | || |    ||   |    ||   ||| | || |||| ||||
Db        537 TATCGGTGTTTCACGTGAAACCCAATTTAATTTATATAATAGCGAATCCTATGCTGGTGG 596

Qy        663 ATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAG 722
               ||  |       ||||||| | ||    ||  |     ||  | ||   ||  |
Db        597 CTTAGTGGAAGTGCTAGGTTATCAAATGCCAAAAGCCCGTGCCGATAACCAACCTAATGC 656

Qy        723 TAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCC 782
              |  |  |  |  || ||                           |  |   ||  || ||
Db        657 TTCGGTTGGTTTAGAACAAG------------------------TGGCGGCAGAAAAGCC 692

Qy        783 TGAGCGTATGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACT 842
              ||| |  ||| || ||||     |         |  | | ||   |     || | ||
Db        693 TGATCTGATGATCTTAATCCATTA---------TCGTGATGAGAGTATTGCAAGAAAATG 743

Qy        843 AGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTT 902
               ||||| ||  |  ||||||| |   |    || || ||||||   |  ||     |
Db        744 GGAAAATGAAGCCTTATGGAAAATTATCCCTGCGGTAAAAAATGGTCAAGTCATCTTAGC 803
```

```
Qy        903 AGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGA 962
                |   | | ||||||||||||| ||||||||| |   | |  | |||| ||||||| |||||
Db        804 TAATGATAATTTATGGGCAAGAGCACGTGGTATTGATGCCGCTGAAGTAATGGCTAAAGA 863

Qy        963 ACTTGTTGAATTATCTAAGAAAGATAGTAAAAAA 996
              | |    || ||    | ||||      ||| ||
Db        864 AGTCCAAGACTTTGTAACGAAATCCGCCAAATAA 897


RESULT 16
US-09-974-300-2066
; Sequence 2066, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
;  APPLICANT: Berka, Randy M.
;  APPLICANT:  Clausen, Ib Groth
;  TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;  TITLE OF INVENTION:  Expression
;  FILE REFERENCE: 10085.500-US
;  CURRENT APPLICATION NUMBER: US/09/974,300
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/680,598
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/279,526
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 8481
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2066
;   LENGTH: 801
;   TYPE: DNA
;   ORGANISM: Bacillus licheniformis
US-09-974-300-2066

  Query Match             14.8%;  Score 149.4;  DB 9;  Length 801;
  Best Local Similarity   51.8%;  Pred. No. 4.8e-16;
  Matches  414;  Conservative    0;  Mismatches  376;  Indels    9;  Gaps    3;

Qy        176 AACGTGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAAC 235
              ||   ||||| |||||||| |    |||| |||||||| |  ||| |   | | |||| |
Db          1 AAAAAGTTGTGGTTCTTGAATTGGGATTTATTGATGCGCTCCTTGATGCGGGTATTAAGC 60

Qy        236 CTGTTGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAA 295
              |||| || || ||||| ||   ||||   ||   |||||   |     | || |  ||||
Db         61 CTGTCGGAATTGCGGACGACGGCAAACCTAAGTTTATTAACGAGAAGGTCAGGGGAAAAA 120

Qy        296 TTGGAAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAAC 355
              |   |   || ||||| || ||  | ||    || || | ||| ||| ||||   |
Db        121 TCAAAGGGTATACTTCAGTCGGTTCGCGCGCCCAGCCAAGCTTTGAAAAAATTGCTTCTT 180

Qy        356 TTAAACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAA 415
              | || || |||||||||||||| |||   |  || ||||  ||  | ||| |  |  |
Db        181 TAAAGCCCGATTTAATTATTGCCGATTCGAGCAGGCACAGCGGCGTCTATGATAAGCTGT 240

Qy        416 ATAAAATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTG 475
                |||||||| || || || |  || ||||  ||  |||  |||||| | || |   |||
Db        241 CGAAAATTGCGCCGACAATCGCGCTCAAAAACTTGAATGCCGATTATCAGGACACGCTTG 300

Qy        476 ATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAG 535
              |||| | |   || ||| | |||||  | || ||||||   | |    | |||   || |
Db        301 ATGCATCTCTTACGATTGCGAAAGCCGTCGGCAAAGAAAGCGCAATGGAGAAAAAATTGG 360

Qy        536 AAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAA 595
                |||||  |  | ||  | || ||||  ||| |  || |    |     ||| |    |
Db        361 CTGAACATAAACAAAAGCTGGATGAATTGAAACAGAAATTCGGCAGCCGGAAACAGAGCA 420

Qy        596 AGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATG 655
                 |  |||  |      |   | ||  |       |     ||  |   |||| ||
Db        421 TTCTTCTGCTCGGGAACACAAATGAAGAAATCACCGTGCGCGATGAA---AACTTTTTCA 477

Qy        656 TTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAG 715
                  |||  | || |   || | || | || | | |   |  | ||    |     ||||
Db        478 CGTCTCAGCTTCTGACGAAAATCGGCTATACATACGGTGTCGGAGACAGCG---GCAAAG 534
```

```
Qy        716 GTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAG 775
              |        |  |   | ||       |  |  ||||||     |||     |   | |
Db        535 GCGATGCCGAAAACGGTGAATCCGTCAATATTAAAATGACGCTCGAACAGCTGCTTGAGA 594

Qy        776 TGAATCCTGAGCGTATGTTCATAATGAC---AAACAAAGCAAGTTCTAACGAACCTTCAC 832
                 |||| ||   |||  || | |||||   | | ||||  |      | |||
Db        595 AAGATCCCGACGTTATTGTCCTGATGACGGGAGAAAAAGATAAAGTCGATGAAGACGGGA 654

Qy        833 TAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTG 892
               || |    | |||||||||||| | ||||| ||  | | ||| || |||||       |
Db        655 AAAGACCGATCGAAAAAGATCCTCTTTGGAAAAAGCTCAGCGCAGTCAAAAACGGCAAGG 714

Qy        893 TTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAA 952
              || ||      ||| |   |   ||| |      |||  |  |         |||| ||
Db        715 TTTATGAAGCCGACAGATTCGCCTGGTCGCTCCGACGCAGCATTGACGGAGCAGATGAGC 774

Qy        953 TGGCAAAAGAACTTGTTGA 971
              ||    | ||| | | | |
Db        775 TGATGGATGAAATCGATCA 793


RESULT 17
US-10-282-122A-9415
; Sequence 9415, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9415
;   LENGTH: 972
;   TYPE: DNA
;   ORGANISM: Bacillus anthracis
US-10-282-122A-9415

  Query Match             14.2%;  Score 143.2;  DB 17;  Length 972;
  Best Local Similarity   55.0%;  Pred. No. 5.7e-15;
  Matches  308;  Conservative    0;  Mismatches  243;  Indels    9;  Gaps    1;

Qy         25 AAAATTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAAT 84
              |||||| | ||| |  |     || | ||| | ||   ||        | |     |||
Db          7 AAAATTCTCAGTATTTTCATAGTAGTTTTTCTATTCGCTGTTGGATGCGGACAGCAAAAA 66

Qy         85 AATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAA 144
               |      ||   |  |||     ||      |||  | |   |   ||| ||  ||||
Db         67 GAGGAGAAAAAAGAAACAAAAGCGGACAATAAAAATCAAGCTATAACAATTAAACACGCT 126

Qy        145 GAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTT 204
              |||||    |||||  ||  |||||   | |||   || ||||| ||||| |    || |
Db        127 GAAGGGGAAACGAAGTTAGATAAACCAGCGAAAAAAGTAGTTGTACTTGAATGGGTATAT 186

Qy        205 GTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAA 264
                 || |  ||| | ||  | | |||| | || || ||||| || ||   ||| ||  |
Db        187 TCAGAAGACTTATTAGCACTTGGTGTTCAGCCAGTAGGGATGGCAGACATTAAGAATTAT 246

Qy        265 AATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGT 324
              |||       | ||  ||  || | |    | |  | |         ||| || ||||||
Db        247 AATAAATGGGTAAATACAAAAACAAAACCGAGTAAAGATGTTGTAGATGTCGGGACACGT 306

Qy        325 AAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAAT 384
               | ||||| |||||||||||||| ||     | |||||||||||||||||  | | |
Db        307 CAACAACCAAACTTAGAAGAAATTAGCCGTTTAAAACCAGATTTAATTATCACAGCTTCA 366

Qy        385 AATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGAT---------T 435
                  |    ||||  ||| | || || ||| |  ||||||| || ||  |         |
Db        367 TTCCGTGGTAAAGCAATTAAAAATGAATTAGAACAAATTGCACCAACAGTTATGTTTGAT 426

Qy        436 GAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCA 495
                |   | |||     | |   | | |  |||||  |  ||  | ||||||  |||| ||
Db        427 CCATCAACAAGCAATAACGATCACTTTGCTGAAATGACAGAAACATTTAAACAAATTGCA 486

Qy        496 AAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATT 555
              |||||  | || |||||||||||||||||||||   ||||  ||    |||||   | |
Db        487 AAAGCAGTTGGAAAAGAAGAAGAAGGTAAAAAAGTATTAGCTGATATGGATAAAGCATTC 546

Qy        556 GAAGAATATAAAAAAGAAAT 575
              |  ||    |||  | ||||
Db        547 GCTGATGCAAAAGCAAAAAT 566

RESULT 18
US-09-815-242-2107/c
; Sequence 2107, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2107
;   LENGTH: 321
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-09-815-242-2107

  Query Match             12.7%;  Score 128;  DB 9;  Length 321;
  Best Local Similarity   67.8%;  Pred. No. 1.7e-12;
  Matches  179;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy        132 AATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCT 191
              ||| ||  | |||  ||| |||   ||| |    ||| | |||||||||||||||||  |
Db        271 AATAAAAGATGAATTAGGAACTGAAAAAATTAAGAAAAATCCTAAACGTGTTGTTGTATT 212

Qy        192 TGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGA 251
               || |||   |||| ||||   ||||  || |||||| | ||||||||||| || || ||
Db        211 AGAATATAGTTTTGCTGATTATTTAGCAGCATTAGATATGAAACCTGTTGGTATTGCAGA 152

Qy        252 TGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTC 311
              ||||  ||  |  ||   ||| |  ||  || |||||||||| |||||   |||    ||
Db        151 TGATGGCAGCACTAAAAATATAACAAAGTCAGTAAGAGATAAGATTGGGGCATATGAATC 92

Qy        312 TGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAAT 371
               || ||| |  |   |||||| ||  | ||||  || |||||| | ||||| ||||| ||
Db         91 GGTTGGATCTAGACCGCAACCGAATATGGAAGTGATAAGTAAATTAAAACCGGATTTGAT 32

Qy        372 TATTGCTGATAATAATAGACACAA 395
               ||||| |||  ||  ||||| ||
Db         31 CATTGCAGATGTTAGCAGACATAA 8

RESULT 19
US-10-282-122A-4605/c
; Sequence 4605, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4605
;   LENGTH: 321
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-10-282-122A-4605

  Query Match             12.7%;  Score 128;  DB 17;  Length 321;
  Best Local Similarity   67.8%;  Pred. No. 1.7e-12;
  Matches  179;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy        132 AATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCT 191
              ||| ||  | |||  ||| |||   ||| |    ||| | |||||||||||||||||  |
Db        271 AATAAAAGATGAATTAGGAACTGAAAAAATTAAGAAAAATCCTAAACGTGTTGTTGTATT 212

Qy        192 TGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGA 251
               || |||   |||| ||||   ||||  || |||||| | ||||||||||| || || ||
Db        211 AGAATATAGTTTTGCTGATTATTTAGCAGCATTAGATATGAAACCTGTTGGTATTGCAGA 152

Qy        252 TGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTC 311
              ||||  ||  |  ||   ||| |  ||  || |||||||||| |||||   |||    ||
Db        151 TGATGGCAGCACTAAAAATATAACAAAGTCAGTAAGAGATAAGATTGGGGCATATGAATC 92

Qy        312 TGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAAT 371
               || ||| |  |   |||||| ||  | ||||  |   ||||| | ||||| ||||| ||
Db         91 GGTTGGATCTAGACCGCAACCGAATATGGAAGTGATAAGTAAATTAAAACCGGATTTGAT 32

Qy        372 TATTGCTGATAATAATAGACACAA 395
               ||||| |||  ||  ||||| ||
Db         31 CATTGCAGATGTTAGCAGACATAA 8


RESULT 20
US-08-781-986A-604
; Sequence 604, Application US/08781986A
; Publication No. US20030054436A1
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  5255
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,986A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Benson, Bob
;      REGISTRATION NUMBER:  30,446
;      REFERENCE/DOCKET NUMBER:  PB248PP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  604:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  2115 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-781-986A-604

  Query Match             12.4%;  Score 125.4;  DB 8;  Length 2115;
  Best Local Similarity   53.2%;  Pred. No. 7.5e-12;
  Matches  289;  Conservative    1;  Mismatches  247;  Indels    6;  Gaps    1;

Qy        458 ATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAG 517
              ||||||||| ||||||||| || ||||||||| |  | |||||  |||| |||||  |||
Db          1 ATTATAATGCAAATATTGAAGCATTTAAAACAGTCGCTAAAGCAGTAGGCAAAGAGAAAG 60

Qy        518 AAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAA 577
              ||||  | || ||  | ||| | || ||||| | | | |  ||   || |||  ||||
Db         61 AAGGCGAGAAGCGTCTGGAAAAGCATGATAAAATATTAGCGGAGATTAGAAAGAAAATTG 120

Qy        578 CTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTC 637
                  |  ||      ||   | |||  |    |      | | | ||||| || ||  |
Db        121 AACAGAGTACGTTAAAATCTGCATTTGCATTCGGTATCTCAAGAGCAGGTATGTTTATTA 180

Qy        638 ATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAA 697
              ||       |  | | | | || |||||| ||| | || | ||| || ||   | | | |
Db        181 ATAATGAAGATACATTTATGGGACAATTCTTAATTAAAATGGGTATTCAACCTGAAGTCA 240

Qy        698 GTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACA 757
                 | :|    ||||    | |  || |          || |||||  || |  | || |
Db        241 MAAAARAMAAAACTACGCATGTTGGTGAACGCAAGGGTGGTCCTTATATATATTTAAATA 300

Qy        758 CTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTT 817
               ||||    |  |  |  | |||||  |   |||| |  ||   ||  ||  |  ||
Db        301 ATGAAGAACTTGCCAATATCAATCCAAAAGTTATGATTTTAGCCACTGACGGAAAAACGG 360

Qy        818 CTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTG 877
                ||  |         ||   || |     |   | | || |||||   |||||| ||||
Db        361 ACAAAAA------TAGAACGAAATTCATTGATCCTGCAGTTTGGAAATCATTAAAAGCTG 414

Qy        878 TGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAA 937
              ||||| || |    ||| ||    | |||||  |    |||   | ||||| | ||  | |
Db        415 TGAAAGATAACAAAGTTTATGACGTTGACCGAAATAAGTGGTTGAAATCAAGGGGGATTA 474

Qy        938 TTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAG 997
              |  |    |||   |||||| ||||  | |   || |  |  | ||||  |   |||||
Db        475 TCGCAAGTGAAAGTATGGCAGAAGATTTAGAAAAAATTGCAGAAAAAGCAAAATAAAAAT 534

Qy        998 ATA 1000
              | |
Db        535 ACA 537


RESULT 21
US-10-329-624-604
; Sequence 604, Application US/10329624
; Publication No. US20040043037A1
;   GENERAL INFORMATION:
;        APPLICANT: Charles Kunsch
;                   Gil H. Choi
;                   Patrick S. Dillon
;                   Craig A. Rosen
;                   Steven C. Barash
;                   Michael R. Fannon
;        TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;        NUMBER OF SEQUENCES: 5256
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;            COMPUTER: HP Vectra 486/33
;            OPERATING SYSTEM: MSDOS version 6.2
;            SOFTWARE: ASCII Text
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/329,624
;            FILING DATE: 27-Dec-2002
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/956,171
;            FILING DATE: October 20, 1997
;            APPLICATION NUMBER: 60/009,861
;            FILING DATE: January 5, 1996
;            APPLICATION NUMBER: 08/781,986
;            FILING DATE: January 3, 1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Mark J. Hyman
;            REGISTRATION NUMBER: 46,789
;            REFERENCE/DOCKET NUMBER: PB248P1D1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (240) 314-1224
;            TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 604:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 2115 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 604:
US-10-329-624-604

  Query Match             12.4%;  Score 125.4;  DB 18;  Length 2115;
  Best Local Similarity   53.2%;  Pred. No. 7.5e-12;
  Matches  289;  Conservative    1;  Mismatches  247;  Indels    6;  Gaps    1;

Qy        458 ATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAG 517
              ||||||||| ||||||||| || ||||||||| |  | |||||  |||| |||||  |||
Db          1 ATTATAATGCAAATATTGAAGCATTTAAAACAGTCGCTAAAGCAGTAGGCAAAGAGAAAG 60

Qy        518 AAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAA 577
              ||||  | || ||  | ||| | || ||||| | | | |  ||   || |||  ||||
Db         61 AAGGCGAGAAGCGTCTGGAAAAGCATGATAAAATATTAGCGGAGATTAGAAAGAAAATTG 120

Qy        578 CTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGGTTTGCTTGCTC 637
                  |  ||      ||   | |||  |    |      | | | ||||| || ||  |
Db        121 AACAGAGTACGTTAAAATCTGCATTTGCATTCGGTATCTCAAGAGCAGGTATGTTTATTA 180

Qy        638 ATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAAAGAAGCATTAA 697
              ||       |  | | | | || |||||| ||| | || | ||| || ||   | | | |
Db        181 ATAATGAAGATACATTTATGGGACAATTCTTAATTAAAATGGGTATTCAACCTGAAGTCA 240

Qy        698 GTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTTACTTACAAATGAACA 757
                 | :|    ||||    | |  || |          || |||||  || |  | || |
Db        241 MAAAARAMAAAACTACGCATGTTGGTGAACGCAAGGGTGGTCCTTATATATATTTAAATA 300

Qy        758 CTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGACAAACAAAGCAAGTT 817
               ||||    |  |  |  | |||||  |   |||| |  ||   ||  ||  |  ||
Db        301 ATGAAGAACTTGCCAATATCAATCCAAAAGTTATGATTTTAGCCACTGACGGAAAAACGG 360

Qy        818 CTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCCTGTATGGAAGAAATTAAACGCTG 877
                ||  |         ||   || |     |   | | || |||||   |||||| ||||
Db        361 ACAAAAA------TAGAACGAAATTCATTGATCCTGCAGTTTGGAAATCATTAAAAGCTG 414

Qy        878 TGAAAAATCAACGTGTTGATATTTTAGACCGTGACTTATGGGCAAGATCACGTGGTTTAA 937
              ||||| || |    ||| ||    | |||||  |    |||   | |||| | ||  | |
Db        415 TGAAAGATAACAAAGTTTATGACGTTGACCGAAATAAGTGGTTGAAATCAAGGGGGATTA 474

Qy        938 TTTCTTCAGAAGAAATGGCAAAAGAACTTGTTGAATTATCTAAGAAAGATAGTAAAAAAG 997
              |  |    |||   |||||| ||||  | |   || |  |  | ||||  |   |||||
Db        475 TCGCAAGTGAAAGTATGGCAGAAGATTTAGAAAAAATTGCAGAAAAAGCAAAATAAAAAT 534

Qy        998 ATA 1000
              | |
Db        535 ACA 537
```

```
RESULT 22
US-10-470-048B-318
; Sequence 318, Application US/10470048B
; Publication No. US20050037444A1
; GENERAL INFORMATION:
;  APPLICANT: MEINKE ET AL.
;  TITLE OF INVENTION: A METHOD FOR IDENTIFICATION, ISOLATION AND PRODUCTION OF
;  TITLE OF INVENTION:  ANTIGENS TO A SPECIFIC PATHOGEN
;  FILE REFERENCE: SONN:035US
;  CURRENT APPLICATION NUMBER: US/10/470,048B
;  CURRENT FILING DATE:  2003-07-25
;  NUMBER OF SEQ ID NOS: 603
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 318
;   LENGTH: 990
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-10-470-048B-318

  Query Match             12.3%;  Score 124;  DB 21;  Length 990;
  Best Local Similarity   50.7%;  Pred. No. 1.1e-11;
  Matches  412;  Conservative    0;  Mismatches  380;  Indels   20;  Gaps    4;

Qy         27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
              ||||  ||   |  |||      |   | |  || |||     |||| |||| |    ||
Db         12 AATTAAAATGCTTGTTGTTACGCTTGCTTTCCTACTTGTTTTAGCAGGATGTAGTGGGAA 71

Qy         87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
              |      |   | |||  | |  |   |   || ||   |  |  ||| || || | |
Db         72 TTCAAATAAACAATCATCTGATAACAAAGATAAGGAAACAACTTCAATTAAACATGCAAT 131

Qy        147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
               ||||| ||  || |        | || || |||||||||   ||  |  |
Db        132 GGGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCAC 191

Qy        207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
              ||| |     ||  |||||| |||||||||||| ||    |  ||       | | ||||
Db        192 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAA 251

Qy        267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
               |  |   |  ||    ||| | || |  |   | |    |    ||||||   |
Db        252 ACCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACC 311

Qy        327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                 |||||||||||| ||||||  |||| | ||||| || |||||| | ||    ||
Db        312 TGCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGT 371

Qy        387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
              |||| |  ||    ||||  |  | |||  |||||| || ||             || ||
Db        372 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAG 418

Qy        447 TTTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAG 505
              |||| |  || |   |  | ||| | |  ||| |   |||   |||    ||||||||||
Db        419 TTTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAG 478

Qy        506 GTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATA 565
              | |||||| |||||| | || |     |  || |  ||||| | ||| | |  | ||
Db        479 GGAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCC 538

Qy        566 AAAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAAT 622
              |||||||   || | | | | ||||| ||  |  |  ||||   ||   || || ||| |
Db        539 AAAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACT 598

Qy        623 CAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
                 ||     || |||| |||| |  | |  || |||  |              ||  |
Db        599 TCCGT---GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 655

Qy        683 TTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTT 742
               |  || |     | || || | |      ||     |  | |  ||  |||||    |
Db        656 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 715

Qy        743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGA 802
```

```
                |  ||   |     |  ||||   |  |   | ||||  |||| | ||| ||  || | |
Db         716 TCCAACTTACATCTAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAA 775

Qy         803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
                |    |  ||| | | || ||  || || ||
Db         776 AATCAGATCCAAATGCGAAAGATGCTGCATTA 807


RESULT 23
US-10-278-946-15
; Sequence 15, Application US/10278946
; Publication No. US20030153733A1
; GENERAL INFORMATION:
;  APPLICANT: Simpson et al.
;  TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
;  FILE REFERENCE: PB461USD1
;  CURRENT APPLICATION NUMBER: US/10/278,946
;  CURRENT FILING DATE:  2002-10-24
;  PRIOR APPLICATION NUMBER: US 09/830,217
;  PRIOR FILING DATE: 2001-04-24
;  PRIOR APPLICATION NUMBER: PCT/US99/06199
;  PRIOR FILING DATE: 1999-03-18
;  PRIOR APPLICATION NUMBER: US 60/078,682
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: US 60/080,296
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: US 60/084,674
;  PRIOR FILING DATE: 1998-05-07
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
;   LENGTH: 999
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-10-278-946-15

  Query Match             12.3%;  Score 124;  DB 16;  Length 999;
  Best Local Similarity   50.7%;  Pred. No. 1.1e-11;
  Matches  412;  Conservative    0;  Mismatches  380;  Indels   20;  Gaps    4;

Qy          27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
               ||||  ||   |  |||      |   | |  || |||     |||| |||| |    ||
Db          18 AATTAAAATGCTTGTTGTTACGCTTGCTTTCCTACTTGTTTTAGCAGGATGTAGTGGGAA 77

Qy          87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
               |      |   | |||  | |  |   |   ||  ||   |  |  ||| || || | |
Db          78 TTCAAATAAACAATCATCTGATAACAAAGATAAGGAAACAACTTCAATTAAACATGCAAT 137

Qy         147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
                ||||| ||  || |        | || || |||||||||   ||  |  |
Db         138 GGGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCAC 197

Qy         207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
               ||| |     ||  |||||| |||||||||||| ||    |  ||       | | ||||
Db         198 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAA 257

Qy         267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
                |  |   |  ||    ||| | || |  |   | |    |    ||||||   |
Db         258 ACCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACC 317

Qy         327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                  |||||||||||| ||||||  |||| | ||||| || |||||| | ||    ||
Db         318 TGCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGT 377

Qy         387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
               |||| |  ||    ||||  |  | |||  |||||| || ||             || ||
Db         378 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAG 424

Qy         447 TTTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAG 505
               |||| |  || |   |  | ||| | |  ||| |   |||   |||    ||||||||||
Db         425 TTTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAG 484

Qy         506 GTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATA 565
```

```
               | |||||| |||||| | || |     |  || |  ||||| | ||| | |  | ||
Db         485 GGAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCC 544

Qy         566 AAAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAAT 622
               |||||||   || | | | | |||||  ||  |  |  ||||   ||   || || ||| |
Db         545 AAAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACT 604

Qy         623 CAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
                  ||     || |||| |||| |  | |  || |||  |              ||  |
Db         605 TCCGT---GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 661

Qy         683 TTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTT 742
                |  || |     | || || | |      ||     |  | |  ||  |||||    |
Db         662 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 721

Qy         743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGA 802
                |  ||  |     |  ||||   |  |   | ||||  |||| | ||| ||  || | |
Db         722 TCCAACTTACATCTAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAA 781

Qy         803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
                |    |  ||| | | || ||  || || ||
Db         782 AATCAGATCCAAATGCGAAAGATGCTGCATTA 813


RESULT 24
US-10-967-189-15
; Sequence 15, Application US/10967189
; Publication No. US20050053995A1
; GENERAL INFORMATION:
;  APPLICANT: Simpson et al.
;  TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
;  FILE REFERENCE: PB461USD2
;  CURRENT APPLICATION NUMBER: US/10/967,189
;  CURRENT FILING DATE:  2004-10-19
;  PRIOR APPLICATION NUMBER: US 10/278,946
;  PRIOR FILING DATE: 2002-10-24
;  PRIOR APPLICATION NUMBER: US 09/830,217
;  PRIOR FILING DATE: 2002-01-15
;  PRIOR APPLICATION NUMBER: PCT/US99/06199
;  PRIOR FILING DATE: 1999-03-18
;  PRIOR APPLICATION NUMBER: US 60/078,682
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: US 60/080,296
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: US 60/084,674
;  PRIOR FILING DATE: 1998-05-07
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
;   LENGTH: 999
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-10-967-189-15

  Query Match             12.3%;  Score 124;  DB 21;  Length 999;
  Best Local Similarity   50.7%;  Pred. No. 1.1e-11;
  Matches  412;  Conservative    0;  Mismatches  380;  Indels   20;  Gaps    4;

Qy          27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
               ||||  ||   |  |||      |   | |  || |||     |||| |||| |    ||
Db          18 AATTAAAATGCTTGTTGTTACGCTTGCTTTCCTACTTGTTTTAGCAGGATGTAGTGGGAA 77

Qy          87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
               |      |   | |||  | |  |   |   ||  ||   |  |  ||| || || | |
Db          78 TTCAAATAAACAATCATCTGATAACAAAGATAAGGAAACAACTTCAATTAAACATGCAAT 137

Qy         147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
                ||||| ||  || |        | || || |||||||||   ||  |  |
Db         138 GGGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCAC 197

Qy         207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
               ||| |     ||  |||||| |||||||||||| ||    |  ||       | | ||||
Db         198 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAA 257
```

```
Qy       267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
              |  |   |  ||    ||| | || |  |   | |    |    ||||||   |
Db       258 ACCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACC 317

Qy       327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                |||||||||||| ||||||  |||| | ||||| || |||||| | ||    ||
Db       318 TGCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGT 377

Qy       387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
             |||| |  ||    ||||  |  | |||  |||||| || ||             || ||
Db       378 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAG 424

Qy       447 TTTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAG 505
             |||| |  || |   |  | ||| | |  ||| |   |||   |||    ||||||||||
Db       425 TTTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAG 484

Qy       506 GTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATA 565
             | |||||| |||||| | || |     |  || |  ||||| | ||| | |  | ||
Db       485 GGAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCC 544

Qy       566 AAAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAAT 622
             |||||||   || | | | | ||||| ||  |  |  ||||   ||   || || ||| |
Db       545 AAAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACT 604

Qy       623 CAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
                ||     || |||| |||| |  | |  || |||  |              ||  |
Db       605 TCCGT---GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 661

Qy       683 TTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTT 742
              |  || |     | || || | |      ||     |  | |  ||  |||||    |
Db       662 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 721

Qy       743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGA 802
              |  ||  |     |  ||||   |  |   | ||||  |||| | ||| ||  || | |
Db       722 TCCAACTTACATCTAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAA 781

Qy       803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
              |    |  ||| | | || ||  || || ||
Db       782 AATCAGATCCAAATGCGAAAGATGCTGCATTA 813

RESULT 25
US-08-781-986A-238
; Sequence 238, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
;   APPLICANT:  Charles Kunsch
;   TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;   NUMBER OF SEQUENCES:  5255
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Human Genome Sciences, Inc.
;     STREET:  9410 Key West Avenue
;     CITY:  Rockville
;     STATE:  Maryland
;     COUNTRY:  USA
;     ZIP:  20850
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;     COMPUTER:  HP Vectra 486/33
;     OPERATING SYSTEM:  MSDOS version 6.2
;     SOFTWARE:  ASCII Text
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/781,986A
;     FILING DATE:
;     CLASSIFICATION:  435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Benson, Bob
;     REGISTRATION NUMBER:  30,446
;     REFERENCE/DOCKET NUMBER:  PB248PP
```

```
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (301) 309-8504
;     TELEFAX:  (301) 309-8512
; INFORMATION FOR SEQ ID NO:  238:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  3775 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  double
;     TOPOLOGY:  linear
US-08-781-986A-238

  Query Match             12.3%;  Score 124;  DB 8;  Length 3775;
  Best Local Similarity   50.7%;  Pred. No. 1.5e-11;
  Matches  412;  Conservative    0;  Mismatches  380;  Indels   20;  Gaps    4;

Qy        27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
             ||||  ||   |  |||      |   | |  || |||     |||| |||| |    ||
Db        60 AATTAAAATGCTTGTTGTTACGCTTGCTTTCCTACTTGTTTTAGCAGGATGTAGTGGGAA 119

Qy        87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
             |      |   | |||  | |  |   |   || ||   |  |  ||| || || | |
Db       120 TTCAAATAAACAATCATCTGATAACAAAGATAAGGAAACAACTTCAATTAAACATGCAAT 179

Qy       147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
              ||||| ||  || |        | || || |||||||||   ||  |  |
Db       180 GGGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCAC 239

Qy       207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
             ||| |     ||  |||||| |||||||||||| ||    |  ||       | | ||||
Db       240 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAA 299

Qy       267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
              |  |   |  ||    ||| | || |  |   | |    |    ||||||   |
Db       300 ACCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACC 359

Qy       327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                |||||||||||| ||||||  |||| | ||||| || |||||| | ||    ||
Db       360 TGCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGT 419

Qy       387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
             |||| |  ||    ||||  |  | |||  |||||| || ||             || ||
Db       420 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAG 466

Qy       447 TTTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAG 505
             |||| |  || |   |  | ||| | |  ||| |   |||   |||    ||||||||||
Db       467 TTTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAG 526

Qy       506 GTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATA 565
             | |||||| |||||| | || |     |  || |  ||||| | ||| | |  | ||
Db       527 GGAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCC 586

Qy       566 AAAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAAT 622
             |||||||   || | | | | ||||| ||  |  |  ||||   ||   || || ||| |
Db       587 AAAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACT 646

Qy       623 CAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
                ||     || |||| |||| |  | |  || |||  |              ||  |
Db       647 TCCGT---GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 703

Qy       683 TTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTT 742
              |  || |     | || || | |      ||     |  | |  ||  |||||    |
Db       704 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 763

Qy       743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGA 802
              |  ||  |     |  ||||   |  |   | ||||  |||| | ||| ||  || | |
Db       764 TCCAACTTACATCTAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAA 823

Qy       803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
              |    |  ||| | | || ||  || || ||
Db       824 AATCAGATCCAAATGCGAAAGATGCTGCATTA 855

RESULT 26
```

```
US-10-329-624-238
; Sequence 238, Application US/10329624
; Publication No. US20040043037A1
;   GENERAL INFORMATION:
;        APPLICANT: Charles Kunsch
;                   Gil H. Choi
;                   Patrick S. Dillon
;                   Craig A. Rosen
;                   Steven C. Barash
;                   Michael R. Fannon
;        TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;        NUMBER OF SEQUENCES: 5256
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/329,624
;             FILING DATE: 27-Dec-2002
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/956,171
;             FILING DATE: October 20, 1997
;             APPLICATION NUMBER: 60/009,861
;             FILING DATE: January 5, 1996
;             APPLICATION NUMBER: 08/781,986
;             FILING DATE: January 3, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mark J. Hyman
;             REGISTRATION NUMBER: 46,789
;             REFERENCE/DOCKET NUMBER: PB248P1D1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (240) 314-1224
;             TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 238:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 3775 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 238:
US-10-329-624-238

  Query Match             12.3%;  Score 124;  DB 18;  Length 3775;
  Best Local Similarity   50.7%;  Pred. No. 1.5e-11;
  Matches  412;  Conservative    0;  Mismatches  380;  Indels   20;  Gaps    4;

Qy           27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
                ||||  ||   |  |||      |   | |  ||  |||    |||| |||| |    ||
Db           60 AATTAAAATGCTTGTTGTTACGCTTGCTTTCCTACTTGTTTTAGCAGGATGTAGTGGGAA 119

Qy           87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
                |      |   | |||  | |  |   |   || ||   |  |  ||| || || | |
Db          120 TTCAAATAAACAATCATCTGATAACAAAGATAAGGAAACAACTTCAATTAAACATGCAAT 179

Qy          147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
                 ||||| ||  ||  |       | || || |||||||||   ||  |  |
Db          180 GGGTACAACTGAAATTAAAGGGAAACCAAAGCGTGTTGTTACGCTATATCAAGGTGCCAC 239

Qy          207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
                ||| |     ||  |||||| |||||||||||| ||    |  ||         | | ||||
Db          240 TGACGTCGCTGTATCTTTAGGTGTTAAACCTGTAGGTGCTGTAGAATCATGGACACAAAA 299

Qy          267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
                 |  |   |  ||    ||| | || |  |   | |    |    ||||||   |
Db          300 ACCGAAATTCGAATACATAAAAAATGATTTAAAAGATACTAAGATTGTAGGTCAAGAACC 359
```

```
Qy          327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                   |||||||||||| ||||||  |||| | ||||| || |||||| | ||    ||
Db          360 TGCACCTAACTTAGAGGAAATCTCTAAATTAAAACCGGACTTAATTGTCGCGTCAAAAGT 419

Qy          387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
                |||| |  ||    ||||  |  | |||  |||||| || ||              || ||
Db          420 TAGAAATGAAAAAGTTTACGATCAATTATCTAAAATCGCACC-------------AACAG 466

Qy          447 TTTCGATGGAGATTATAATGAAAAT-ATTGATGCTTTTAAAACAATTTCAAAAGCTTTAG 505
                |||| |  || |   |  | ||| | |  ||| |   |||   |||    ||||||||||
Db          467 TTTCTACTGATACAGTTTTCAAATTCAAAGATACAACTAAGTTAATGGGGAAAGCTTTAG 526

Qy          506 GTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATA 565
                | |||||| |||||| | || |     |  || |  ||||| | ||| | |  | ||
Db          527 GGAAAGAAAAAGAAGCTGAAGATTTACTTAAAAAGTACGATGATAAAGTAGCTGCATTCC 586

Qy          566 AAAAAGA---AATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAAT 622
                |||||||   || | | | | ||||| ||  |  |  ||||   ||   || || ||| |
Db          587 AAAAAGATGCAAAAGCAAAGTATAAAGATGCATGGCCATTGAAAGCTTCAGTTGTTAACT 646

Qy          623 CAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTT 682
                   ||     || |||| |||| |  | |  || |||  |              ||  |
Db          647 TCCGT---GCTGATCATACAAGAATTTATGCTGGTGGATATGCTGGTGAAATCTTAAATG 703

Qy          683 TTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAAAGGACCTT 742
                 |  || |     | || || | |      ||     |  | |  ||  |||||    |
Db          704 ATTTAGGATTCAAACGTAATAAAGACTTACAAAAACAAGTTGATAATGGTAAAGATATTA 763

Qy          743 ACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTTCATAATGA 802
                 |  ||  |     |  ||||   |  |   | ||||  |||| | ||| ||  || | |
Db          764 TCCAACTTACATCTAAAGAAAGCATTCCATTAATGAACGCTGATCATATTTTTGTAGTAA 823

Qy          803 CAAACAAAGCAAGTTCTAACGAACCTTCACTA 834
                 |    |  ||| | | || ||  || || ||
Db          824 AATCAGATCCAAATGCGAAAGATGCTGCATTA 855


RESULT 27
US-10-282-122A-10009
; Sequence 10009, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10009
;   LENGTH: 796
;   TYPE: DNA
;   ORGANISM: Bacillus anthracis
US-10-282-122A-10009

  Query Match             11.2%;  Score 113.2;  DB 17;  Length 796;
  Best Local Similarity   57.6%;  Pred. No. 7e-10;
  Matches  227;  Conservative    0;  Mismatches  158;  Indels    9;  Gaps    1;

Qy        191 TTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGG 250
              |||| |    || |   || |  ||| | ||  | | |||| | || || ||||| || |
Db          2 TTGAATGGGTATATTCAGAAGACTTATTAGCACTTGGTGTTCAGCCAGTAGGGATGGCAG 61

Qy        251 ATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTT 310
              |   ||| ||  | |||       | ||  ||  || | |    | |  | |
Db         62 ACATTAAGAATTATAATAAATGGGTAAATACAAAAACAAAACCGAGTAAAGATGTTGTAG 121

Qy        311 CTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAA 370
               ||| || ||||||| | ||||| |||||||||||||| ||     | |||||||||||||
Db        122 ATGTCGGGACACGTCAACAACCAAACTTAGAAGAAATTAGCCGTTTAAAACCAGATTTAA 181

Qy        371 TTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTA 430
              ||||  | | |       |    ||||  ||| | || || ||| | |  ||||||| || |
Db        182 TTATCACAGCTTCATTCCGTGGTAAAGCAATTAAAAATGAATTAGAACAAATTGCACCAA 241

Qy        431 CGAT---------TGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTT 481
              |  |         |  |   | |||     | |   | | |  |||||  |  ||  | |
Db        242 CAGTTATGTTTGATCCATCAACAAGCAATAACGATCACTTTGCTGAAATGACAGAAACAT 301

Qy        482 TTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAAC 541
              |||||  |||| ||||||| | || |||||||||||||||||||||   ||||  ||
Db        302 TTAAACAAATTGCAAAAGCAGTTGGAAAAGAAGAAGAAGGTAAAAAAGTATTAGCTGATA 361

Qy        542 ACGATAAGAAAATTGAAGAATATAAAAAAGAAAT 575
                |||||   | | |  ||    |||  | ||||
Db        362 TGGATAAAGCATTCGCTGATGCAAAAGCAAAAAT 395

RESULT 28
US-10-282-122A-19870
; Sequence 19870, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
```

```
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19870
;   LENGTH: 900
;   TYPE: DNA
;   ORGANISM: Enterobacter cloacae
US-10-282-122A-19870

  Query Match             10.2%;  Score 103.2;  DB 17;  Length 900;
  Best Local Similarity   52.9%;  Pred. No. 3.6e-08;
  Matches  222;  Conservative    0;  Mismatches  198;  Indels    0;  Gaps    0;

Qy        137 AGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGT 196
              || ||||| | || ||    | | |   ||||   ||  ||||  |||| || || ||
Db         74 AGGACGAACACGGCACGTTTACACTCGATAAAACGCCACAACGGATTGTGGTGCTGGAAC 133

Qy        197 ATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
                || || |  |||||| | |  ||  | || || |  ||  | || || || || ||||
Db        134 TCTCGTTCGCCGATGCGCTGGCCGCCGTGGACGTCAGCCCGATCGGTATTGCCGACGATA 193

Qy        257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
              || |   ||| || ||  |      | |  | |   |  |   |   |      || || |
Db        194 ACGATGCAAAACGCATCCTGCCCGAAGTGCGTGCGCACCTGAAACCGTGGCAGTCCGTCG 253

Qy        317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
              |||| ||   ||| || | | | ||||  ||   |   || |||||||| | || ||||
Db        254 GAACGCGCGCGCAGCCGAGCCTGGAAGCCATTGCTGCCCTGAAACCAGACCTGATCATTG 313

Qy        377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
              | || |  | | | ||    || ||||| |  | |||  |  |||| || ||     |
Db        314 CCGACAGCAGTCGCCATGCGGGGATTTACACTGCCTTGCAGCAAATCGCGCCGGTACTGC 373

Qy        437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
                || ||      | | | |   ||   ||||||| |  |  |       |  ||
Db        374 TGCTTAAGTCCCGCAACGAAACCTACGCTGAAAATTTGCACTCAGCGGCTATCATCGGCG 433

Qy        497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
              |||   ||||||||| ||  |||     |    ||  | ||| ||||  |  |||| || |
Db        434 AAGTGGTAGGTAAAAAACGAGAGATGCAGGCGCGTCTGGAACAACATAAAGAGAAGATGG 493

RESULT 29
US-09-815-242-6372
; Sequence 6372, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6372
;   LENGTH: 909
;   TYPE: DNA
;   ORGANISM: Escherichia coli
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(909)
US-09-815-242-6372

  Query Match             9.4%;  Score 95.2;  DB 9;  Length 909;
  Best Local Similarity   51.7%;  Pred. No. 8.4e-07;
  Matches  217;  Conservative    0;  Mismatches  203;  Indels    0;  Gaps    0;

Qy        137 AGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGT 196
              || ||||| | || ||    | | |    |||   ||  ||||  |||| || || ||
Db         80 AGGACGAACACGGCACGTTTACACTCGAAAAAACGCCACAACGGATTGTGGTGCTGGAAC 139

Qy        197 ATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
                || || |  |||||| | |  ||  | || || |  ||  | || || || || ||||
Db        140 TCTCGTTCGCCGATGCGCTGGCCGCCGTGGACGTCATCCCGATCGGTATTGCCGACGATA 199

Qy        257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
              || |   ||| || || |       | |  | |   |  |   |   |     || || |
Db        200 ACGATGCAAAACGCATCCTGCCCGAAGTGCGTGCGCACCTGAAACCGTGGCAGTCCGTCG 259

Qy        317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
              |||| ||   ||| || | | | ||||  ||       || ||||||||  | || ||||
Db        260 GAACGCGCGCGCAGCCGAGCCTGGAAGCCATTGCCGCTCTGAAACCAGACCTGATCATTG 319

Qy        377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
              | || |  | | | ||    ||  |||| |  | |||  |  |||| || ||     |
Db        320 CCGACAGCAGTCGCCATGCGGGGGTTTACATCGCCTTGCAGCAAATCGCGCCGGTACTGC 379

Qy        437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
                || ||      | | | |   ||   ||||||| |  |  ||      |  ||
Db        380 TGCTTAAGTCCCGCAACGAAACCTACGCTGAAAATTTGCAATCTGCGGCTATCATCGGCG 439

Qy        497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
              ||    | |||||| |   |||     |   |||  | ||| ||||  |  |||  || |
Db        440 AAATGGTGGGTAAAAAGCGAGAGATGCAGGCACGTCTGGAACAACATAAAGAGAGGATGG 499

RESULT 30
US-10-282-122A-20621
; Sequence 20621, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
```

```
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20621
;   LENGTH: 909
;   TYPE: DNA
;   ORGANISM: Escherichia coli
US-10-282-122A-20621

  Query Match             9.4%;  Score 95.2;  DB 17;  Length 909;
  Best Local Similarity   51.7%;  Pred. No. 8.4e-07;
  Matches  217;  Conservative    0;  Mismatches  203;  Indels    0;  Gaps    0;

Qy        137 AGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGT 196
              || ||||| | || ||    | | |    |||   ||  ||||  |||| || || ||
Db         80 AGGACGAACACGGCACGTTTACACTCGAAAAAACGCCACAACGGATTGTGGTGCTGGAAC 139

Qy        197 ATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
                || || |  |||||| | |  ||  | || || |  ||  | || || || || ||||
Db        140 TCTCGTTCGCCGATGCGCTGGCCGCCGTGGACGTCATCCCGATCGGTATTGCCGACGATA 199

Qy        257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
              || |   ||| || || |       | |  | |   |  |   |   |     || || |
Db        200 ACGATGCAAAACGCATCCTGCCCGAAGTGCGTGCGCACCTGAAACCGTGGCAGTCCGTCG 259

Qy        317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
              |||| ||   ||| || | | | ||||  ||       || ||||||||  | || ||||
Db        260 GAACGCGCGCGCAGCCGAGCCTGGAAGCCATTGCCGCTCTGAAACCAGACCTGATCATTG 319

Qy        377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
              | || |  | | | ||    ||  |||| |  | |||  |  |||| || ||     |
Db        320 CCGACAGCAGTCGCCATGCGGGGGTTTACATCGCCTTGCAGCAAATCGCGCCGGTACTGC 379

Qy        437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
                || ||      | | | |   ||   ||||||| |  |  ||      |  ||
Db        380 TGCTTAAGTCCCGCAACGAAACCTACGCTGAAAATTTGCAATCTGCGGCTATCATCGGCG 439

Qy        497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
              ||    | |||||| |   |||     |   |||  | ||| ||||  |  |||  || |
Db        440 AAATGGTGGGTAAAAAGCGAGAGATGCAGGCACGTCTGGAACAACATAAAGAGAGGATGG 499

RESULT 31
US-10-893-671-14/c
```

```
; Sequence 14, Application US/10893671
; Publication No. US20050064527A1
; GENERAL INFORMATION:
;  APPLICANT: Levy, Stuart, et. al.
;  TITLE OF INVENTION: NIMR COMPOSITIONS AND THEIR METHODS OF USE
;  FILE REFERENCE: PKZ-043
;  CURRENT APPLICATION NUMBER: US/10/893,671
;  CURRENT FILING DATE:  2004-07-15
;  PRIOR APPLICATION NUMBER: US/09/801,563
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: 60/188,362
;  PRIOR FILING DATE: 2000-03-10
;  NUMBER OF SEQ ID NOS: 98
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 10244
;   TYPE: DNA
;   ORGANISM: Escherichia coli
US-10-893-671-14

  Query Match              9.4%;  Score 95.2;  DB 21;  Length 10244;
  Best Local Similarity   51.7%;  Pred. No. 1.6e-06;
  Matches  217;  Conservative    0;  Mismatches  203;  Indels    0;  Gaps    0;

Qy         137 AGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGT 196
               || ||||| | || ||    | | |    |||   ||  ||||  |||| || || ||
Db        7731 AGGACGAACACGGCACGTTTACACTCGAAAAAACGCCACAACGGATTGTGGTGCTGGAAC 7672

Qy         197 ATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATA 256
                 || || |  |||||| | |  ||  | || || |  ||  | || || || || ||||
Db        7671 TCTCGTTCGCCGATGCGCTGGCCGCCGTGGACGTCATCCCGATCGGTATTGCCGACGATA 7612

Qy         257 ACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAG 316
               || |   ||| || ||  |      | | | |   |  |   |   |     || || |
Db        7611 ACGATGCAAAACGCATCCTGCCCGAAGTGCGTGCGCACCTGAAACCGTGGCAGTCCGTCG 7552

Qy         317 GAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTG 376
               |||| ||   ||| || | | | ||||  ||       || ||||||||  | || ||||
Db        7551 GAACGCGCGCGCAGCCGAGCCTGGAAGCCATTGCCGCTCTGAAACCAGACCTGATCATTG 7492

Qy         377 CTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTG 436
               | || |  | | | ||    ||  |||| |  | |||  |  |||| || ||     |
Db        7491 CCGACAGCAGTCGCCATGCGGGGGTTTACATCGCCTTGCAGCAAATCGCGCCGGTACTGC 7432

Qy         437 AACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAA 496
                 || ||      | | | |   ||   ||||||| |  |  ||      |  ||
Db        7431 TGCTTAAGTCCCGCAACGAAACCTACGCTGAAAATTTGCAATCTGCGGCTATCATCGGCG 7372

Qy         497 AAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTG 556
               ||    | |||||| |   |||     |   |||  | ||| ||||  |  |||  || |
Db        7371 AAATGGTGGGTAAAAAGCGAGAGATGCAGGCACGTCTGGAACAACATAAAGAGAGGATGG 7312

RESULT 32
US-10-282-122A-17083
; Sequence 17083, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
```

```
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17083
;   LENGTH: 963
;   TYPE: DNA
;   ORGANISM: Clostridium difficile
US-10-282-122A-17083

  Query Match              7.8%;  Score 79;  DB 17;  Length 963;
  Best Local Similarity   47.1%;  Pred. No. 0.00049;
  Matches  312;  Conservative    0;  Mismatches  345;  Indels    6;  Gaps    2;

Qy          27 AATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAATAA 86
               |||  ||   | ||| |   |  |     |  | |  |    ||||  | |||||| |||
Db          18 AATAGTAGCAGCAATAGCTATTATAGGATTAGTGACAGTGTTTGCATTAGGTGGAAGTAA 77

Qy          87 TAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGAAGA 146
                |             ||    |||    |||||  ||     |  | || |  ||  |
Db          78 AAAGAATGAATCAAAAACATCAGAAAATTCAAATAATACCATAAAGATAACTCATAACTT 137

Qy         147 AGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATTTGT 206
                ||    ||  | |||     | | ||||||   || || ||  |||| |||||   | |
Db         138 GGGAGAAACAGATGTAAAATTAAATCCTAAAAAAGTAGTAGTATTTGATTATTCGGCTTT 197

Qy         207 TGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAAAAA 266
                ||| |  | | ||| |||| |||   | |||      |||  |    |   |||   |
Db         198 AGATACAATGGATGCATTAGGTGT---AGCTGAGAATTTAGTAGGACTTCCAAAAGCGAG 254

Qy         267 TCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACGTAA 326
               |              |  || | ||||   |   || || ||   | |||||      ||
Db         255 TTTACCAGCATCTTTAGAAAAATATAAGGATAAGAAGTATACAGATTTAGGAGGCTTAAA 314

Qy         327 GCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAATAA 386
                 ||||  | |||||||  || |        || ||||||||||||||     ||   |
Db         315 AGAACCAGATTTAGAAGGTATAAAGTCGGCAAATCCAGATTTAATTATAATAAATGGAAG 374

Qy         387 TAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAAAAG 446
                  | |  |   |  | |  ||   | ||| | |    |   | |  | |  |   |||
Db         375 ACAAGAAGATTTTTATGAGCAATTGTCAAAAATAGCACCAACAATAAGTACAAGTAAAGA 434

Qy         447 TTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCAAAAGCTTTAGG 506
               |    |      | |   |      ||   ||  | |  | | |||  |||||   || ||
Db         435 TGATAAGAAGTACTTAGAATCTGTTAAAAATAATATAGACAAAATAGCAAAAATATTTGG 494

Qy         507 TAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGAATATAA 566
                  ||||||| |||  ||  ||   ||     ||   || |||||||| |||   |  ||
Db         495 AGTAGAAGAAAAAGCAAATCAAGAATTTAGTAAAATTGAAAAGAAAATAGAAACTTTGAA 554

Qy         567 AAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGCTGCTAAATCAGG 626
                ||| || ||||    |||||||||  ||| |  ||  ||  | | |    | ||
Db         555 TAAAAAAGTAACA---GATAAAAATTTAAATGCTTTAACTATAATGGTAAATGAAGGTAA 611
```

```
Qy          627 TTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCAACTAGGTTTTAA 686
                |||   ||    |       | ||     || ||  |||    | |    | || ||| |
Db          612 TTTAAGTGTATTTGGTGAAGAATCAAGATTTAGTATATTATACAATAGTTTTGGATTTGA 671

Qy          687 AGA 689
                | |
Db          672 AAA 674


RESULT 33
US-10-282-122A-32601
; Sequence 32601, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32601
;   LENGTH: 1026
;   TYPE: DNA
;   ORGANISM: Proteus mirabilis
US-10-282-122A-32601

  Query Match              7.8%;  Score 78.8;  DB 17;  Length 1026;
  Best Local Similarity   48.9%;  Pred. No. 0.00054;
  Matches  271;  Conservative    0;  Mismatches  277;  Indels    6;  Gaps    2;

Qy           24 AAAAATTTTAAGTGTAATTGGCTTATTGTTTGTTTTAATTGCAACTGCAGCATGTGGAAA 83
                ||||   |||||   | |   ||||||| |   ||   | |  |  ||||  ||||  ||
Db           78 AAAATCATTAAGCCCACTGTTCTTATTGCTGTCTTCTCTGGTTATCGCAGGTTGTGATAA 137

Qy           84 TAATAGTTCAAGTAACTCAAGTAAAGAGTCATCAAAAGATGGAGTTGAAATCAAGCACGA 143
                          |      |||   ||    ||   ||| |     |    ||| | || |
Db          138 CGCACAAGATACCTCAACAACGCAATCCACAGAGAAACAAACTCTCACTATCGAACATGC 197

Qy          144 AGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTGTTCTTGAGTATTCATT 203
                  |||| || ||  |  | |||   |||||  ||   || |||||  |  |      |
Db          198 TCAAGGCACCACTGAGATCCCTGCTCACCCACAAAAAGTGGTTGTGATGAACATGGAAAC 257

Qy          204 TGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAGCGGATGATAACAAAAA 263
                  |||||    | | |||| | |  ||      ||||||  || |  |    |||    |
Db          258 ACTTGATATTGTTGATGCTCTTGGCGTACCTGTTGTTGGCCTACCACAAACAAACGTCCA 317

Qy          264 AAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGTAGGAACACG 323
                    | ||   ||  |     |      ||        | ||||     | |||    |
Db          318 CCTACCTAAATTTCTAGAGAAATACACCAACGAAAATGAGTACATCAACGAAGGTGGCCT 377

Qy          324 TAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTATTGCTGATAA 383
                     |||| ||||  ||| || ||||||    |   || ||| | ||| |  | | ||
Db          378 GTTTGAACCAAACTACGAAAAACTCAGTACTACTGCCCCTGATCTGATTTTAACGGGTAG 437

Qy          384 TAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGATTGAACTGAA 443
                |  |   | |   | |   |||    | |||| | |||| || ||  | |||    || |
Db          438 TCGTGCTCGC---GATGCATATGCCAAATTAAGTGAAATCGCACCATCCATTTCTATGGA 494

Qy          444 AAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTCA---AAAGC 500
                 |    |     |   | || ||| | | |   ||    || ||| | || |    ||
Db          495 TATCGACAGCACACGCTTTATTGACAGTCTCACTGAGCGTACAACGACTTTAGGTCAAAT 554

Qy          501 TTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAATTGAAGA 560
                ||| |||||||||||| ||| |||||||   ||||  ||   | |||  ||||||||
Db          555 TTTTGGTAAAGAAGAACAAGCTAAAAAATTATTAGCTGACTTCAATAGCAAAATTGATAC 614

Qy          561 ATATAAAAAAGAAA 574
                    |||  | |||
Db          615 GGTAAAAGCAAAAA 628


RESULT 34
US-09-960-352-5558/c
; Sequence 5558, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 5558
;   LENGTH: 431
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 24-LIB3057-024-Q1-K1-F7
US-09-960-352-5558

  Query Match              7.7%;  Score 78;  DB 9;  Length 431;
  Best Local Similarity   51.4%;  Pred. No. 0.00059;
  Matches  180;  Conservative    0;  Mismatches  170;  Indels    0;  Gaps    0;

Qy          254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
                |||| |||||| |    |  | ||||  |  || | |||||     |||| |    |
Db          424 ATAAAAAAAAATAAAAAAAAAATAAAATAAAAAAAAATAAATAAATAAAAAATTAATAAA 365

Qy          314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
                 |  || |   ||  |    ||   | || |||| |  |||  ||||  | |  |||| |
Db          364 AATAAATAATAAAAAATATAAAAAAAAAAAAAAATAAAAAAAAATAAAATATAAATAATAA 305

Qy          374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                    | | ||  |   | | |||  ||   | ||| |  |||| |||| |      |  |
Db          304 AAAATAAAAAACAATAAAAAAAAAAATAAAAAAAAAATAAATAAAAAAAAAATAAAAAAAAAA 245

Qy          434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                || ||   ||||      |   | |  | ||  |||| ||  |   |  ||||| ||
Db          244 TTAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAATAATAAAAAAAAAAA 185
```

```
Qy        494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
               ||||     |   ||| |  || ||   ||| ||     | || || |  |||  ||||
Db        184 TAAAATAAAAAAAAAAAAATAAAAAAAAAAAATAAAAAAAAAAAAAAAATAATATAAAAA 125

Qy        554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTG 603
               | || |  | |||| | ||| ||  |   | || ||  ||||  ||| |
Db        124 ATAAAAATAAAAAAATAAAAAAAAAAAAATAAAATAAATAAAATTTATAG 75


RESULT 35
US-10-021-323-6774
; Sequence 6774, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
;  APPLICANT: Deikman, Jill
;  APPLICANT:  Feng, Paul C.C.
;  APPLICANT:  Fincher, Karen L.
;  APPLICANT:  Ziegler, Todd E.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(52274)B
;  CURRENT APPLICATION NUMBER: US/10/021,323
;  CURRENT FILING DATE:  2001-12-12
;  PRIOR APPLICATION NUMBER: US 60/255, 619
;  PRIOR FILING DATE: 2000-12-14
;  NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 6774
;   LENGTH: 478
;   TYPE: DNA
;   ORGANISM: Gossypium hirsutum
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(478)
;   OTHER INFORMATION: unsure at all n locations
;   OTHER INFORMATION: Clone ID: LIB3828-010-Q1-N6-E1
US-10-021-323-6774

  Query Match             7.7%;  Score 77.8;  DB 19;  Length 478;
  Best Local Similarity   49.3%;  Pred. No. 0.00065;
  Matches  202;  Conservative    0;  Mismatches  208;  Indels    0;  Gaps    0;

Qy        180 TGTTGTTGTTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGT 239
              | || || || ||   | ||  ||| ||  |   ||  ||  |||   | || || |  | |
Db          6 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAATTTTT 65

Qy        240 TGGGATAGCGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGG 299
              |   | |    |  | || ||||||||    |  |  |||  |  || | | ||||
Db         66 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 125

Qy        300 AAAATACACTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAA 359
              |||| | |       |  || |   ||  ||   ||   | || |||  |  |||   ||
Db        126 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 185

Qy        360 ACCAGATTTAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAA 419
              |  | |   ||  |      | || || | | | |||   |   | ||| |   ||| ||
Db        186 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 245

Qy        420 AATTGCTCCTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGC 479
              ||        |     |||  |||||      |   | |  | ||  |||| |   |
Db        246 AAAAAGGAAAAAATAGGAAAAGAAAAAAAGGAAAATAAAAAAAAAAAAAAAAAAGAAAAA 305

Qy        480 TTTTAAAACAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGA 539
                  |||| ||    ||||     | | ||| || || ||    |||||     | ||||
Db        306 GGAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAATGNAAAAAGAAAAAAAAGA 365

Qy        540 ACACGATAAGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAA 589
              | | ||   ||||| |||| || | |||||  | | ||  |   | ||||
Db        366 AAAGGAGGGGAAAAATGAAAAAAAAAAAAAGAATAAAAAGAAAAAAAAAA 415


RESULT 36
US-10-282-122A-33236
; Sequence 33236, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33236
;   LENGTH: 891
;   TYPE: DNA
;   ORGANISM: Pseudomonas syringae
US-10-282-122A-33236

  Query Match             7.5%;  Score 75.8;  DB 17;  Length 891;
  Best Local Similarity   47.7%;  Pred. No. 0.0017;
  Matches  221;  Conservative    0;  Mismatches  242;  Indels    0;  Gaps    0;

Qy        128 TTGAAATCAAGCACGAAGAAGGTACTACGAAAGTACCTAAACACCCTAAACGTGTTGTTG 187
              | || ||| |  |||   |    |    | |  | ||  |   ||| || || || || |
Db         53 TCGACATCGATGACGGCCAGCACAAGGTGCATCTGCCAGACACCCCCAAGCGCGTGGTGG 112

Qy        188 TTCTTGAGTATTCATTTGTTGATGCGTTAGTTGCTTTAGATGTTAAACCTGTTGGGATAG 247
              | || || | ||| ||| | || | | | | | |  | |  || | ||| || ||    |
Db        113 TACTGGAATTTTCGTTTCTCGACGGGCTTGCTTCGGTTGGCGTGACACCGGTCGGCGCCG 172

Qy        248 CGGATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACA 307
              | ||||||  | |    | |||| |  |        |  |  |     | ||  |||
Db        173 CAGATGATGGCGACGCCAGTCGTGTACTGCCCAAGGTGCGCAAGGCCGTGGGTGAATGGC 232

Qy        308 CTTCTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATT 367
                || || ||    ||   |||||| ||  | ||||  |||      || || || ||
Db        233 AGTCGGTGGGGCTGCGCTCGCAACCCAATATCGAAGTGATCGCACGGCTCAAGCCGGACC 292

Qy        368 TAATTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTC 427
              | || || |||||      | | ||  | | | ||||||| ||| |    |   | ||||
Db        293 TGATCATCGCTGACCTCGGTCGTCATCAGGCTCTTTATAACGACCTGGCCAGCCTGGCTC 352

Qy        428 CTACGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAA 487
```

```
               | ||| |     ||        ||  |  || ||| | |  |   |  |  ||
Db         353 CGACGCTGATGTTGCCGTCACGCGGCGAGGACTATCAGGGCAGCCTGAAATCTGCCGGGC 412

Qy         488 CAATTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATA 547
                 ||    |  ||  | || ||||     ||     |    ||  | | |||| ||  |
Db         413 TGATCGGCATGGCGCTGGGCAAAGGCCCGGAGATGCAGGCGCGAATCGCAGAAAACCGTC 472

Qy         548 AGAAAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAA 590
               || |  |  ||     |    |  | ||  ||   || |  ||
Db         473 AGCACCTGAAAACCGTTGCCGAGCAGATTCCTGCCGACAGCAA 515


RESULT 37
US-10-425-115-115706/c
; Sequence 115706, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 115706
;   LENGTH: 1121
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(1121)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_37006C.1
US-10-425-115-115706

  Query Match             7.5%;  Score 75.6;  DB 20;  Length 1121;
  Best Local Similarity   44.7%;  Pred. No. 0.0019;
  Matches  288;  Conservative    0;  Mismatches  356;  Indels    0;  Gaps    0;

Qy         255 TAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTGT 314
               ||| || |||||    |  |  |||| |  || |  |||||    |||| | |      |
Db        1089 TAAAAATAAAAAGAAAACAAAAAAACAAAAAAAAATTAAAAAAAAAAAAAAAAAGAAAAT 1030

Qy         315 AGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTAT 374
               |  |  |   ||       |||  | || |||  |   ||    ||  | |   ||  |
Db        1029 AAAATGAAAAAAAAGTAAAAACGAAAAAAAAAAAAAATAAAAACAAAAAAAACAAANAAA 970

Qy         375 TGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGAT 434
                    | || || | | |||||   |   | ||| |   |||  | |    |   |  |
Db         969 AAAAAAAAAAAAAAAAAACAAAAAAAGAAAAAAATAAAAAAAACATAAAAATAAAAAAAG 910

Qy         435 TGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTTC 494
                 |||  ||||   |  |   | |  |  |  ||||||   |       || |  |
Db         909 AAAACAAAAAAAAATGAAAAAAAAAAAAGAAAAAAATAAAAAAAAAAAAAATAACAAAAA 850

Qy         495 AAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAAT 554
               ||||   |     ||| || || ||   ||||||     | ||  | |  | || ||||
Db         849 AAAAAAATGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAGAAAAAAAAAAA 790

Qy         555 TGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAGTAGC 614
                 || || | |||||| ||| ||  |   | |||||  ||||   |        |  |
Db         789 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 730

Qy         615 TGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAAGTCA 674
                   |||  |         | | |   ||  ||   | |        ||     ||   |
Db         729 AAAAAAAAAAAAAAAAAAAGATAAAATAAAAAAAAATAAAAAAATAAAAAAAAAAAAAAA 670

Qy         675 ACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATCTTAA 734
               |  |    ||||| || |   |    |  |   |  |||     ||  |  |||  |||
Db         669 AAAAAAAATTAAAAAAAAAAATAAAAGTAATAAAAAAAAAACAAAAAAAATATATAATAA 610

Qy         735 AGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTATGTT 794
               |     |    || |||  || | | | | |  |     |   |    | |    || |
Db         609 AAACAATATTATAAAAAAAAAAAATAACAATAAAAAAACAAAAACAAATAAAATAATATA 550

Qy         795 CATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAGATCC 854
                | ||  | ||| |||  || |   |    |    ||| |||| ||  |  |||| || |
Db         549 AAAAACCATAAAAAAAAAAAATATAATAACAAAAACACAAAAAATAAATAAAATATNC 490

Qy         855 TGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATA 898
                 ||   || | || |||       ||| ||  |     |||||
Db         489 ACTACATAATAGATCAAAAAAAACCAAACATACAAAACCTGATA 446


RESULT 38
US-10-425-115-120013/c
; Sequence 120013, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 120013
;   LENGTH: 1062
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(1062)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_40935C.1
US-10-425-115-120013

  Query Match             7.4%;  Score 74.2;  DB 20;  Length 1062;
  Best Local Similarity   51.0%;  Pred. No. 0.0033;
  Matches  175;  Conservative    0;  Mismatches  168;  Indels    0;  Gaps    0;

Qy         254 ATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
               | || ||||||||    |  || |||  |  || | | ||||    |||| |
Db         919 AAAAAAAAAAAAAAAAAAAAATAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAGAAAAGA 860

Qy         314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
                |  || ||  ||  ||   ||   | || |||  |  |||   | |  | |   |||||
Db         859 AAAAAAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAACAAATTA 800

Qy         374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                   | | || || | | | |||   |   | ||| |   ||| ||||        |  |
Db         799 AAAATAAAAAAAAAAAAAAAAAAAACAAAAAAAAATATAAAAAAAAAAAAAAAAAAAAAA 740

Qy         434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                  ||   ||||      |   | |  | ||  |||| ||  |    |  |||| |||
Db         739 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAATAAAAAAAAATAA 680

Qy         494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
                ||||     | | ||| || || |    | ||||     | || || |  | || ||||
Db         679 AAAAAAAAAAAAGAAAAAAAAAAAATAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 620

Qy         554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAA 596
                  || || | |||||| ||| ||  ||  | |||||  ||||
Db         619 AAAAATAAAAAAAAAAAAAAAAAAAAATTAAAAAAAAAAAAAA 577


RESULT 39
US-10-424-599-102083
```

```
; Sequence 102083, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 102083
;   LENGTH: 1214
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(1214)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_63198C.1
US-10-424-599-102083

  Query Match             7.3%;  Score 74;  DB 18;  Length 1214;
  Best Local Similarity   48.7%;  Pred. No. 0.0037;
  Matches  167;  Conservative    0;  Mismatches  176;  Indels    0;  Gaps    0;

Qy        254 ATAACAAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTTCTG 313
              |||| ||||||||    |  |  |||| | ||| |   || |    |||| | |
Db        860 ATAAAAAAAAAAAAAAAAAAAAAANAAAACAAATAAAAATAAATAAAAAAAAAANAAAAAAATAA 919

Qy        314 TAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAATTA 373
               |  || |  ||   ||   |    | || |||  |  |||   |||  | |   ||  |
Db        920 AAAAAAGAAATACAAAAAAAAATAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAA 979

Qy        374 TTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTACGA 433
                    | || || |   | ||    |   | ||| |   ||| ||||        |  |
Db        980 AANAAAAAAAAAAAAAAAANAANAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1039

Qy        434 TTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAATTT 493
                 ||   ||||      |   | |  | ||  |||| |   |       |||| ||
Db       1040 AAAAAAAAAAAAAAAAAAAAANNAAAAAANANAAAAAAAAAAAAANAANAAAAAAAAAAAAAAANAA 1099

Qy        494 CAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGAAAA 553
               ||||     |   ||| || || ||   ||||||     | || || |  | || ||||
Db       1100 AAAAAANAANAANAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAA 1159

Qy        554 TTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAA 596
                 || || | |||||| ||| ||  |   | |||||  ||||
Db       1160 NNNAAAAAAAAAAAAAAAANAAAAAAAAAAAAANAAAAAAAAAAAAA 1202

RESULT 40
US-10-425-115-172717/c
; Sequence 172717, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 172717
;   LENGTH: 1243
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(1243)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_89100C.1
US-10-425-115-172717

  Query Match             7.3%;  Score 74;  DB 20;  Length 1243;
  Best Local Similarity   43.7%;  Pred. No. 0.0037;
  Matches  330;  Conservative    0;  Mismatches  422;  Indels    3;  Gaps    1;

Qy        251 ATGATAACAAAAAAAATCGTATTATTAAACCATTAAGAGATAAAATTGGAAAATACACTT 310
              |  | || |||||||     |  |  |||  |  || | |  |||    |||| | |
Db       1230 AANAAAAAAAAAAAAAANNAAAAAAAACAAAAAAANAAAAAAAAANAAAAAAAAAAAAAAAAAA 1171

Qy        311 CTGTAGGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAAACCAGATTTAA 370
                     ||||   |   ||   ||     || |||  |   ||   |||  | |   ||
Db       1170 AAAACAAAACAAAAACAAAAAAAAAAAAACAAAAAAAAAAAANNAAAAAAAAAAAAAAAAANAA 1111

Qy        371 TTATTGCTGATAATAATAGACACAAAGGTATTTATAAAGACTTAAATAAAATTGCTCCTA 430
                |      | || || |   |||||   |   | ||| |   ||| ||||        |
Db       1110 AAAAAAAAAAAAAAAAAAAANAACAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAA 1051

Qy        431 CGATTGAACTGAAAAGTTTCGATGGAGATTATAATGAAAATATTGATGCTTTTAAAACAA 490
                |   ||   ||||      |   | |  | ||  |||| |   |  |   ||||| ||
Db       1050 ATAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAAATAAAAAAA 991

Qy        491 TTTCAAAAGCTTTAGGTAAAGAAGAAGAAGGTAAAAAACGCTTAGAAGAACACGATAAGA 550
               |  ||||     |   |||  | || ||   ||||||       || || |  | || |
Db        990 ATAAAAAAAAAAAAAAAAAAAATAAAAAAAAAACAAAAAAAAAAAATAAAAAATAAAAAAAAA 931

Qy        551 AAATTGAAGAATATAAAAAAGAAATAACTATGGATAAAAATCAAAAGGTATTGCCTGCAG 610
              |||   || || | |||| | ||| ||  |   | |||||| ||||   |        |
Db        930 AAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAATAAAAATAAAAAAAAAAAAA 871

Qy        611 TAGCTGCTAAATCAGGTTTGCTTGCTCATCCAAGCAACTCTTATGTTGGTCAATTCCTAA 670
              ||      |||  |  ||         |   ||  ||   | |   |    |||    ||
Db        870 TAAAAAAAAAAAAAAAAATTAAAAATAAAAAATAAAAAAAAATAAAAATAAAAAATATAGAA 811

Qy        671 GTCAACTAGGTTTTAAAGAAGCATTAAGTGATGATGTTACTAAAGGTTTAAGTAAGTATC 730
                 || ||  | | |||  |  || ||   |  |    |  ||||    ||  ||  |
Db        810 AAAAAATAAATATAAAAATAAAATAAAAAAAAAAAAAAAAAAAAAGTAAAAATAAAAAAAA 751

Qy        731 TTAAAGGACCTTACTTACAAATGAACACTGAAACTTTATCTCAAGTGAATCCTGAGCGTA 790
                  |  |   ||   | | | ||| |   | |    |  | ||   ||          |
Db        750 AATTAAAAAAATAAAGAAACAAGAAAAAAAAATAAAAAAAATTAAANAAAAAAAAACAAAAA 691

Qy        791 TGTTCATAATGACAAACAAAGCAAGTTCTAACGAACCTTCACTAAAAGAACTAGAAAAAG 850
                |    ||| | ||| ||   ||     ||  ||     |  || | ||  | ||||
Db        690 AAT---NAATTAAAAATAATANAAAAAAAAAAATAAAAAAAAAAAAAGAAAAAAAANAAAAGA 634

Qy        851 ATCCTGTATGGAAGAAATTAAACGCTGTGAAAAATCAACGTGTTGATATTTTAGACCGTG 910
              || || |||  ||  ||  ||| |     | |||  ||   |     |    | |  |
Db        633 ATTCTATATATAAATAAAAAAAAAGGAANAATAAAAAAAAAAGGAATTAAAAAAAAAAAAGAA 574

Qy        911 ACTTATGGGCAAGATCACGTGGTTTAATTTCTTCAGAAGAAATGGCAAAAGAACTTGTTG 970
              ||        ||    |        |||    | | || | |    |||| |     ||
Db        573 ACAAGAAAAAAAATTAAAAAAAAAAGAAATACACTTATAAAATAAAAAAAAAATATTANATTA 514

Qy        971 AATTATCTAAGAAAGATAGTAAAAAAGATAATAAG 1005
              ||| ||  |  | | |||  |||||  | |  |||
Db        513 AATAATAAATAATANATAACAAAAACAAAATAAAG 479

Search completed: August 26, 2005, 12:35:16
Job time : 2126 secs
```